AGGREGATION-DIFFUSION TO CONSTRAINED INTERACTION: MINIMIZERS & GRADIENT FLOWS IN THE SLOW DIFFUSION LIMIT

KATY CRAIG AND IHSAN TOPALOGLU

ABSTRACT. Inspired by recent work on minimizers and gradient flows of constrained interaction energies, we prove that these energies arise as the slow diffusion limit of well-known aggregation-diffusion energies. We show that minimizers of aggregation-diffusion energies converge to a minimizer of the constrained interaction energy and gradient flows converge to a gradient flow. Our results apply to a range of interaction potentials, including singular attractive and repulsive-attractive power-law potentials. In the process of obtaining the slow diffusion limit, we also extend the well-posedness theory for aggregation-diffusion equations and Wasserstein gradient flows to admit a wide range of nonconvex interaction potentials. We conclude by applying our results to develop a numerical method for constrained interaction energies, which we use to investigate open questions on set valued minimizers.

CONTENTS

1. INTRODUCTION

Nonlocal interactions arise throughout the natural world, from collective dynamics in biological swarms to vortex motion in superconductors and gravitational interactions among stars. In each case, agents experience pairwise attractive or repulsive forces, and these pairwise interactions are often coupled with additional repulsive effects, such as diffusion or a height constraint, which penalize accumulations. The simplest mathematical model for nonlocal interactions and diffusion is the *aggregation-diffusion equation*,

$$\partial_t \rho - \nabla \cdot ((\nabla K * \rho)\rho) = \Delta \rho^m, \quad K : \mathbb{R}^d \to \mathbb{R}, \quad m \geqslant 1,$$

where the *interaction potential* K governs the pairwise interactions and the *diffusion exponent* $m \geqslant 1$ controls the strength at which diffusion is felt at different heights of the density ρ.

Likewise, nonlocal interactions coupled with a height constraint can be heuristically modeled by the *constrained aggregation equation*

$$\begin{cases} \partial_t \rho - \nabla \cdot ((\nabla K * \rho)\rho) = 0 & \text{if } \rho < 1, \\ \rho \leqslant 1 \text{ always}, \end{cases}$$

where, again, $K : \mathbb{R}^d \to \mathbb{R}$ is the interaction potential. We note that this equation is merely a heuristic partial differential equation, as we do not specify the sense in which the height constraint $\rho \leqslant 1$ is enforced. We provide a rigorous formulation below.

Both the aggregation-diffusion equation and constrained aggregation equation have gradient flow structures with respect to the *2-Wasserstein metric*. The aggregation-diffusion equation is formally the gradient flow of the sum of an interaction energy and Rényi entropy

$$(1.1) \quad \mathsf{E}_m(\rho) = \begin{cases} \dfrac{1}{2} \iint K(x-y)\, \rho(x)\rho(y)\, dx\, dy + \dfrac{1}{m-1} \int \rho(x)^m\, dx & \text{if } \rho \in L^m(\mathbb{R}^d), \\ +\infty & \text{otherwise}, \end{cases}$$

and the constrained aggregation equation can be rigorously posed as the gradient flow of the constrained interaction energy

$$(1.2) \quad \mathsf{E}_\infty(\rho) = \begin{cases} \dfrac{1}{2} \iint K(x-y)\, \rho(x)\rho(y)\, dx\, dy & \text{if } \rho \in L^\infty(\mathbb{R}^d) \text{ and } \|\rho\|_\infty \leqslant 1, \\ +\infty & \text{otherwise}. \end{cases}$$

Over the past fifteen years, there has been significant work on aggregation-diffusion equations, analyzing dynamics of solutions, asymptotic behavior, and minimizers of the energy E_m [8, 9, 12, 17, 19, 22, 26, 28, 29, 31, 34, 37, 49, 70, 71]. The vast majority of the literature has considered one of two choices of interaction potential: either purely attractive power-laws or repulsive-attractive power-laws,

$$(1.3) \quad K(x) = |x|^p/p \quad \text{or} \quad K(x) = |x|^q/q - |x|^p/p \quad \text{for} \quad 2-d \leqslant p < q \leqslant 2, \quad q > 0,$$

with the convention that $|x|^0/0 = \log(|x|)$. For the purely attractive case, the literature has largely studied the competition between the attraction parameter p and the diffusion exponent m, along with the effects this competition has on properties such as global existence of solutions or finite time blowup; see [10, 14–16, 18, 19, 23, 24, 32, 67]. For the repulsive-attractive case, the requirement $p < q$ ensures that the nonlocal interactions are repulsive at short length scales and attractive at long length scales. This competition between short-range and long-range effects leads to rich pattern formation in both the steady states of solutions and the minimizers of the corresponding energy E_m; see [6, 7, 13, 33, 35, 40, 43, 45, 65].

More recently, several works have also considered the constrained aggregation equation and minimizers of the constrained interaction energy E_∞. Minimizers of E_∞ are directly related to a shape optimization problem introduced by Burchard, Choksi, and the second author [20]: given a repulsive-attractive power-law interaction potential K, as in equation (1.3),

$$(1.4) \quad \text{minimize} \quad \mathsf{E}(\Omega) = \frac{1}{2} \int_\Omega \int_\Omega K(x-y)\, dx dy \quad \text{over sets } \Omega \subseteq \mathbb{R}^d \text{ of volume } M.$$

Competition between the attraction parameter q and the repulsion parameter p in the definition of K determines existence, nonexistence, and qualitative properties of minimizers, providing a counterpoint to the well-studied nonlocal isoperimetric problem. (See [41] for a survey.) Burchard, Choksi, and the second author showed that the shape optimization problem admits a solution if and only if the constrained interaction energy E_∞ admits a *set valued minimizer*, i.e., a minimizer ρ that is a characteristic function of a set Ω, $\rho = \chi_\Omega$. Furthermore, they

proved that for attraction $q = 2$ and repulsion $-d < p < 0$, there are critical values of the mass $M_1 \leqslant M_2$ so that set valued minimizers with mass M exist for $M \geqslant M_2$ and do not exist for $M \leqslant M_1$. Subsequently, Frank and Lieb extended this result to $q > 0$ and $p = 2 - d$ and proved that there are also critical values of the mass that separate the *liquid* and *solid* phases of minimizers of E_∞: if $M < M_1^*$, then minimizers of E_∞ satisfy $|\{\rho = 1\}| = 0$ (liquid), and if $M > M_2^*$, then $|\{\rho = 1\}| = M$ (solid) [46] . On one hand, it is known that

$$(1.5) \qquad\qquad M_1^* \leqslant M_1 \leqslant M_2 \leqslant M_2^*,$$

and for Newtonian repulsion and quadratic attraction ($p = 2 - d$, $q = 2$, $d > 2$), all four values equal 1. On the other hand, Lopes provided an explicit example for which $M_1^* < M_2^*$ [55]. In general, it remains unknown for which values of p and q strict inequality holds in any of the three inequalities in (1.5), as well as how the values of the critical masses depend on p and q.

Concurrently with this work on minimizers of the constrained interaction energy E_∞, Kim, Yao, and the first author studied gradient flows of E_∞, which formally solve the constrained aggregation equation [44]. This work was inspired by the vast literature on height constrained problems, which arise in both models of crowd motion and tumor growth (see, e.g., [38, 39, 48, 56, 57, 59, 61]). In the case of a purely attractive Newtonian interaction potential (equation (1.3) with $p = 2 - d$), they characterized gradient flows of E_∞ with set valued initial data in terms of a Hele-Shaw type free boundary problem. A key element of their proof was that, formally, gradient flows of E_m converge to gradient flows of E_∞ as $m \to +\infty$. Indeed, Alexander, Kim, and Yao had proved the analogous results for *drift* diffusion equations in previous work [1]. However, in the case of *aggregation* diffusion equations, rigorous analysis of this limit was not considered, due to the lack of convexity of the interaction potential K.

The objective of the present work is to prove that, indeed, minimizers and gradient flows of E_m do converge to minimizers and gradient flows E_∞ in the *slow diffusion limit* as $m \to +\infty$. We consider measures with a fixed mass $M > 0$, and without loss of generality, we rescale so that $M = 1$. For a general class of interaction potentials K, including both attractive and repulsive-attractive power-law potentials (1.3), we prove that minimizers of E_m converge to a minimizer of E_∞ (up to a subsequence and translations) and gradient flows of E_m converge to a gradient flow of E_∞ (up to a subsequence) in the weak-* topology of probability measures. The latter result extends the famous Mesa Problem for the porous medium equation to include a singular nonlocal interaction term (see e.g. [21]).

In the process of proving these results, we also rigorously prove the equivalence between solutions of aggregation-diffusion equations and gradient flows of the energies E_m. Likewise, we extend the well-posedness theory for aggregation-diffusion equations to include singular repulsive-attractive power-law potentials, thereby filling a gap in the existing theory. Finally, we succeed in characterizing the minimal subdifferential of E_∞ along the gradient flow, a key quantity in the study of gradient flows, which was identified formally in previous works on constrained energies [56, 57]. We believe that one of our main contributions is the extension of the theory of Wasserstein gradient flows to energies, such as E_m and E_∞, that satisfy neither the classical λ-convexity assumption of Ambrosio, Gigli, and Savaré [2] nor the more recent ω-convexity assumption [3, 4, 27, 29, 42].

Finally, we apply these theoretical results to develop a numerical method for gradient flows and minimizers of the constrained interaction energy E_∞. We use Carrillo, Patacchini, and the first author's *blob method for diffusion* (see [35]) to simulate gradient flows and minimizers of E_m for m large, thereby approximating the corresponding gradient flows and minimizers of E_∞. While there exist other numerical methods for constrained problems—such as Liu, Wang, and Zhou's method for purely attractive Newtonian interactions [53] and several Eulerian methods

based on the JKO scheme [25,38,47,62]—our particle method is unique in its ability to resolve the nonlocal interaction term for a range of interaction potentials K. As the primary goal of the present work is theoretical analysis of the slow diffusion limit, we restrict our numerical study to one dimension, though our method naturally extends to all dimensions $d \geqslant 1$.

We conclude with several numerical simulations that shed light on open questions for minimizers of the constrained interaction energy. These numerical results indicate that the critical values of the mass M_1 and M_2 that separate nonexistence and existence of set valued minimizers of E_∞ are in fact equal, and we explore how $M_1 = M_2$ depends on the attraction and repulsion parameters q and p. We also observe that, for $p = 1$, the critical masses M_1^* and M_2^* that separate the liquid and solid phases are in general not equal, except for $q = 2$, so that the existence of an intermediate phase is indeed the generic behavior for minimizers of the constrained interaction energy.

We now describe the assumptions we impose on the interaction potentials K and then provide a precise statement of our main results. We conclude the introduction with an outline of our approach and a brief summary of our notation.

1.1. **Assumptions on Interaction Potentials.** We impose the following assumptions on the interaction potential K and diffusion exponent m. To ensure lower semicontinuity of the energies E_m and E_∞ with respect to weak-* convergence of measures, we suppose that $m \geqslant m_0$, where m_0 and the interaction potential satisfy the following condition.

(LSC) $K : \mathbb{R}^d \to [-\infty, \infty]$ is even, locally integrable, and $K = K_a + K_b$ for two lower semicontinuous functions K_a and K_b, where K_a is bounded below and $K_b \in L^{r,\infty}(\mathbb{R}^d)$, for $r \in (1, +\infty)$; the lower bound on the diffusion exponent satisfies $m_0 > 1 + 1/r$.

Remark 1.1 (Diffusion Dominated Regime)*.* In the case of an attractive power-law interaction potential, $K(x) = |x|^p/p$ for $-d < p < 0$, hypothesis (LSC) is equivalent to the requirement that we are in the diffusion dominated regime, $m_0 > 1 - p/d$ (cf. [10,31,67]). More generally, for repulsive-attractive power-law interaction potentials $K(x) = |x|^q/q - |x|^p/p$ with $-d < p < q$, hypothesis (LSC) merely requires that $m_0 > 1$. (See Proposition 3.3.)

In order to establish the existence of compactly supported minimizers and prove that minimizers of E_m converge to a minimizer of E_∞ as $m \to +\infty$, we impose the following assumptions on the regularity and growth of the interaction potential.

(ATT) Either K is purely attractive and approaches some constant $\ell \in \mathbb{R}$ at infinity or K grows to infinity at infinity. Namely, either

(i) $\lim_{|x| \to +\infty} K(x) = \ell$, and $\ell - K \in L^p(\mathbb{R}^d \setminus B)$ for some $1 \leqslant p < \infty$;
$K \in C^1(\mathbb{R}^d \setminus \{0\})$, $\partial_{|x|} K \geqslant 0$, and for all $|x| > 1$, $\partial_{|x|} K \leqslant C$; or,

(ii) $\lim_{|x| \to +\infty} K(x) = +\infty$.

To prove that gradient flows of E_m converge to a gradient flow of E_∞, we impose the following assumptions on the growth, regularity, and stability of $K * \rho$ and $\nabla K * \rho$ for all $\rho \in \mathcal{P}_2(\mathbb{R}^d) \cap L^m(\mathbb{R}^d)$, where $\mathcal{P}_2(\mathbb{R}^d)$ denotes the set of probability measures with finite second moment, $M_2(\rho) = \int |x|^2 d\rho(x) < +\infty$. We assume that there exists a constant $C \geqslant 1$ and a continuous, nondecreasing, concave function $\psi : [0, +\infty) \to [0, +\infty)$ with $\psi(0) = 0$ so that for all $m \geqslant m_0$ and $\rho, \mu, \nu \in \mathcal{P}_2(\mathbb{R}^d) \cap L^m(\mathbb{R}^d)$,

(GF1) $\|\nabla K * \rho\|_{L^2(\nu)} \leqslant C(1 + \|\rho\|_m + M_2(\rho)^{1/2} + M_2(\nu)^{1/2})$;

(GF2) $K * \rho \in C^1(\mathbb{R}^d)$ and $\left| \nabla K * \rho(x) - \nabla K * \rho(y) \right|^2 \leqslant C(1 + \|\rho\|_m^2) \psi(|x - y|^2)$;

(GF3) $\|\nabla K * (\rho - \nu)\|_{L^2(\mu)} \leqslant C\Big(1 + \|\rho\|_m + \|\nu\|_m + \|\mu\|_m\Big) \psi(d_{W_{2-\epsilon}}(\rho, \nu))$ for some $\epsilon \in (0, 1)$.

Remark 1.2 (differentiability vs convexity). Hypothesis (GF2) is weaker than the analogous hypotheses in previous work [42, 45], since in the present context we merely require differentiability of E_m instead of convexity (or ω-convexity) of E_m.

Hypotheses (LSC), (ATT), and (GF1)-(GF3) are satisfied by the attractive and repulsive-attractive power law potentials described in the introduction (1.3); see Theorem 3.1.

1.2. **Main Results.** Our first main result establishes the convergence of energy minimizers.

Theorem 1.3 (minimizers weak-* converge to minimizer). *Suppose K satisfies hypotheses (LSC) and (ATT). Then for any sequence $\rho_m \in \mathcal{P}_2(\mathbb{R}^d)$ of minimizers of E_m, there exists $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ so that, up to a subsequence and translations, $\rho_m \overset{*}{\rightharpoonup} \rho$ in $\mathcal{P}(\mathbb{R}^d)$ and ρ minimizes E_∞.*

Remark 1.4 (existence of minimizers of E_m). We use hypothesis (ATT) to conclude existence of minimizers of E_m. If one were able to obtain existence by other means, hypothesis (LSC) is sufficient to conclude the $m \to +\infty$ limit.

For attractive or repulsive-attractive power-law potentials, we adapt the arguments by Rein [63] and Frank and Lieb [46], respectively, to prove that minimizers of the energies E_m are compactly supported *uniformly* in m.

Theorem 1.5 (uniform bound on support). *Let ρ_m be a minimizer of the energy E_m. For*

$$(1.6) \quad K(x) = |x|^p/p \quad with \ -d < p < 0 \quad or \quad K(x) = \frac{1}{q}|x|^q - \frac{1}{p}|x|^p \quad with \ -d < p < 0 < q$$

there exists $R > 0$ so that $\operatorname{supp} \rho_m \subset B_R(0)$ for all $m > 1$ sufficiently large.

As a consequence of the previous two theorems, we obtain the convergence of minimizers of E_m to a minimizer of E_∞ in the stronger 2-Wasserstein distance.

Corollary 1.6 (minimizers converge to minimizer). *For interaction potentials K of the form (1.6), any sequence of minimizers of E_m converges, up to a subsequence and translations, to a minimizer of E_∞ in the 2-Wasserstein metric.*

We next turn our attention to gradient flows of the energies E_m and E_∞. We begin by showing that, for m sufficiently large, gradient flows of E_m exist and solve the aggregation-diffusion equation, for all initial data in the domain of the energy $D(\mathsf{E}_m) = \{\rho \in \mathcal{P}_2(\mathbb{R}^d) : \mathsf{E}_m(\rho) < +\infty\}$.

Theorem 1.7 (well-posedness of gradient flows). *Suppose K satisfies hypotheses (LSC) and (GF1)–(GF3) and $m \in [m_0, +\infty]$.*

  (i) *For all $\rho_m^{(0)} \in D(\mathsf{E}_m)$, the gradient flow of E_m with initial data $\rho_m^{(0)}$ exists.*
  (ii) *If the modulus $\psi(s)$ in (GF2)–(GF3) satisfies $\psi(s) \geqslant s$ and $\int_0^1 (s\psi(s))^{-1/2} ds = +\infty$, then the gradient flow is unique.*
  (iii) *For $m < +\infty$ and $\rho_m^{(0)} \in D(\mathsf{E}_m)$, $\rho_m(t)$ is a gradient flows of E_m if and only if it solves the aggregation-diffusion equation in the duality with $C_c^\infty(\mathbb{R}^d \times [0, T])$,*

$$\partial_t \rho_m + \nabla \cdot ((\nabla K * \rho_m)\rho_m) = \Delta \rho_m^m, \quad \rho_m(0) = \rho_m^{(0)}$$

Remark 1.8 (existence). In the particular case that K is a singular attractive power-law potential, $K(x) = |x|^p/p$ for $2 - d \leqslant p \leqslant 0$, and the diffusion exponent is sufficiently large, $m \geqslant m_0 > \max\{d/(d + p - 1), 1\}$, the previous theorem extends the range of initial data ρ_0 for which it is known that solutions to the aggregation-diffusion equation exist from $\rho_0 \in L^\infty(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^d)$ to $\rho_0 \in D(\mathsf{E}_m)$; see [10, 11, 67]. In Proposition 4.10, we also strengthen the energy dissipation *inequality* from these previous works to an energy dissipation *identity*.

To our knowledge, the previous theorem provides the first existence results for aggregation-diffusion equations with singular repulsive-attractive power-law potentials of the form $K(x) = |x|^q/q - |x|^p/p$, $2 - d \leqslant p < q \leqslant 2$, provided that $m \geqslant m_0 > \max\{d/(d + p - 1), 1\}$. This complements recent work by Carrillo and Wang, which studied global boundedness of solutions, under the assumption that solutions exist locally in time [33].

Remark 1.9 (uniqueness). If $K(x) = |x|^p/p$ or $K(x) = |x|^q/q - |x|^p/p$ for $2 - d \leqslant p < q \leqslant 2$ and $m_0 \geqslant d/(p + d - 2)$, then we may take $\psi(s) = s|\log(s)|$ for s near zero in hypotheses (GF2)–(GF3); see [42, Proposition 4.4]. Consequently, the gradient flow of E_m is unique for $m \geqslant m_0 \geqslant d/(p + d - 2)$ and the gradient flow of E_∞ is unique.

We apply this result to show that, up to a subsequence, gradient flows of E_m with well-prepared initial data converge to a gradient flow of E_∞ as $m \to +\infty$.

Theorem 1.10 (subsequence of gradient flows converges to gradient flow)**.** *Suppose K satisfies hypotheses (LSC) and (GF1)–(GF3). Let $\rho_m(t)$ be a gradient flow of E_m. Suppose that the initial data $\rho_m^{(0)}$ is well-prepared:* $\sup_m M_2(\rho_m^{(0)}) < +\infty$ *and for some $\rho^{(0)} \in D(\mathsf{E}_\infty)$*

$$\rho_m^{(0)} \overset{*}{\rightharpoonup} \rho^{(0)} \text{ weak-* in } \mathcal{P}(\mathbb{R}^d) \text{ and } \lim_{m \to \infty} \mathsf{E}_m(\rho_m^{(0)}) = \mathsf{E}_\infty(\rho^{(0)}).$$

Then ρ_m has a weak- convergent subsequence so that $\rho_m(t) \overset{*}{\rightharpoonup} \rho(t)$ for almost every $t \geqslant 0$, and $\rho(t)$ is a gradient flow of E_∞ with initial data $\rho^{(0)}$. Furthermore, as $m \to \infty$,*

$$\mathsf{E}_m(\rho_m(t)) \to \mathsf{E}_\infty(\rho(t)) \text{ for all } t \geqslant 0;$$

$$|\partial \mathsf{E}_m|(\rho_m) \to |\partial \mathsf{E}_\infty|(\rho) \text{ and } |\rho_m'|_{d_W} \to |\rho'|_{d_W} \text{ in } L^2_{\text{loc}}(0, +\infty).$$

Finally, for almost every $t \geqslant 0$, there exists

(1.7) $\sigma(t) \in H^1(\mathbb{R}^d)$ *satisfying $\sigma(t) \geqslant 0$ and $\sigma(t) = 0$ almost everywhere on $\{\rho(t) < 1\}$,*

so that, up to a subsequence, $\rho_m^m(t) \rightharpoonup \sigma(t)$ in $L^2(\mathbb{R}^d)$ and $\nabla K * \rho(t) + \nabla \sigma(t)/\rho(t)$ is the element of $\partial E_\infty(\rho(t))$ with minimal $L^2(\rho(t))$ norm.

Remark 1.11 (minimal subdifferential of E_∞). A byproduct of our result on the convergence of gradient flows is that we are able to characterize the minimal element of the subdifferential of E_∞ along the gradient flow. This result can be easily extended to allow a λ-convex drift potential $V(x)$ in the energies E_m and E_∞ by adding term of the form ∂V to the subdifferential. This makes rigorous formal characterizations of the subdifferential from previous works [56, 57].

For attractive or repulsive-attractive power-law potentials, we may use uniqueness of the gradient flow of E_∞ (see Remark 1.9) to immediately obtain a stronger convergence result.

Corollary 1.12 (gradient flows converge to gradient flow)**.** *Given an interaction potential*

$$K(x) = |x|^p/p \quad or \quad K(x) = |x|^q/q - |x|^p/p \quad for \quad 2 - d \leqslant p < q \leqslant 2,$$

consider gradient flows $\rho_m(t)$ of E_m with well-prepared initial data: $\sup_m M_2(\rho_m^{(0)}) < +\infty$ *and for some $\rho^{(0)} \in D(\mathsf{E}_\infty)$*

$$\rho_m^{(0)} \overset{*}{\rightharpoonup} \rho^{(0)} \text{ weak-* in } \mathcal{P}(\mathbb{R}^d) \text{ and } \lim_{m \to \infty} \mathsf{E}_m(\rho_m^{(0)}) = \mathsf{E}_\infty(\rho^{(0)}).$$

Then $\rho_m(t) \overset{*}{\rightharpoonup} \rho(t)$ for almost every $t \geqslant 0$, where $\rho(t)$ is the unique gradient flow of E_∞ with initial data $\rho^{(0)}$.

1.3. **Outline and Notation.** The remainder of the paper is organized as follows. In section 2, we recall fundamental results on the Wasserstein metric, gradient flows, and Γ-convergence. In section 3, we prove that attractive and repulsive-attractive power-law potentials satisfy our main hypotheses (LSC), (ATT), and (GF1)-(GF3) (Theorem 3.1). In section 4, we prove that the energies E_m and E_∞ are lower semicontinuous with respect to weak-* convergence of measures (Proposition 4.3) and bounded below, uniformly in m (Proposition 4.1). We then characterize the minimal element of the subdifferential of E_m (Proposition 4.7), identify an element of subdifferential of E_∞ (Proposition 4.9), and prove well-posedness of gradient flows (Theorem 1.7). In section 5, we prove our main results on convergence of minimizers (Theorem 1.3) and the uniform bound on the support of minimizers (Theorem 1.5). In section 6, we prove our main result on convergence of gradient flows (Theorem 1.10). Finally, in section 7, we apply these theoretical results to develop a numerical method for simulating gradient flows and minimizers of E_∞, which we use to explore the open questions about minimizers of E_∞ described in the introduction.

We conclude by briefly reviewing our notation. When a probability measure $\rho \in \mathcal{P}(\mathbb{R}^d)$ is absolutely continuous with respect to Lebesgue measure, $\rho \ll \mathcal{L}^d$, we commit a mild abuse of notation and denote both the measure and its density by ρ, $d\rho = \rho(x)dx$. Differentials in integrals will likewise be written either as $d\rho(x)$ or $\rho(x)dx$, depending on the context. Norms with respect to Lebesgue measure will be denoted by single subscripts (e.g., $\|\cdot\|_p$) whereas L^p-norms with respect to a measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ will be explicitly marked (e.g., $\|\cdot\|_{L^p(\mu)}$). We denote convergence with respect to the weak-* topology by $\overset{*}{\rightharpoonup}$. For measures that depend on time $\mu(t) \in \mathcal{P}(\mathbb{R}^d)$, we commit a mild abuse of notation and identify

$$(\mu_t)_{t \in (0,1)} \sim \int_0^1 \delta_t \otimes \mu_t dt \text{ so that } \iint f(t,x)d\mu = \int_0^1 \int_{\mathbb{R}^d} f(t,x)d\mu_t(x)dt.$$

We let χ_Ω denote the characteristic function on a set $\Omega \subseteq \mathbb{R}^d$ and Ω^c denote the complement of Ω. We allow all constants $C > 0$ to change from line to line.

2. PRELIMINARIES

2.1. **The Wasserstein Metric.** For $b \in [1,2]$, we consider measures belonging to the space

$$\mathcal{P}_b(\mathbb{R}^d) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) \colon \int |x|^b d\mu(x) < +\infty \right\}$$

of probability measures with finite bth moments. We endow this space with the b-Wasserstein metric, which we recall briefly now. For further background, we refer the reader to the books by Ambrosio, Gigli and Savaré [2] and Villani [69].

The b-Wasserstein distance between $\mu, \nu \in \mathcal{P}_b(\mathbb{R}^d)$ is given by

$$(2.1) \qquad d_{W_b}(\mu, \nu) := \left(\min \left\{ \iint |x-y|^b \, d\gamma(x,y) \colon \gamma \in \mathcal{C}(\mu, \nu) \right\} \right)^{1/b},$$

where $\mathcal{C}(\mu, \nu)$ is the set of transport plans between μ and ν,

$$\mathcal{C}(\mu, \nu) := \left\{ \gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \colon (\pi_1)_\# \gamma = \rho \quad \text{and} \quad (\pi_2)_\# \gamma = \nu \right\}.$$

Here π_1, π_2 denote the projections $\pi_1(x,y) = x$ and $\pi_2(x,y) = y$. For $i = 1, 2$, $(\pi_i)_\# \gamma$ denotes the pushforward of γ defined by $(\pi_i)_\# \gamma(U) := \gamma(\pi_i^{-1}(U))$ for any measurable set $U \subset \mathbb{R}^d$. By

Hölder's inequality for the probability measure $\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$, we have

$$(2.2) \qquad\qquad d_{W_b}(\mu, \nu) \leqslant d_{W_a}(\mu, \nu), \qquad \text{for all } b \leqslant a.$$

For any $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, the minimization problem (2.1) admits a solution: there exists an optimal transport plan $\gamma_0 \in \mathcal{C}_0(\mu, \nu)$ so that

$$d_{W_b}(\mu, \nu) = \left(\iint |x - y|^b \, d\gamma_0(x, y) \right)^{1/b}.$$

Furthermore, if $b \geqslant 1$ and ν is absolutely continuous with respect to Lebesgue measure,

$$\nu \in \mathcal{P}_{b,ac}(\mathbb{R}^d) := \left\{ \rho \in \mathcal{P}_b(\mathbb{R}^d) \colon \rho \ll \mathcal{L}^d \right\},$$

then there exists an optimal transport plan γ_0 that is given by the product of the identity map $\mathrm{Id}(x) = x$ and a Borel measurable function $\mathbf{t}_\nu^\mu : \mathbb{R}^d \to \mathbb{R}^d$, i.e., $\gamma_0 = (\mathrm{Id} \times \mathbf{t}_\nu^\rho)_{\#}\nu$ (cf. [2, Theorem 6.2.10], [5, Theorem 7.1]). The function $\mathbf{t}_\nu^\mu$ is an *optimal transport map* from ν to μ.

Along with these characterizations of optimal transport plans, for all $b \in [1, 2]$, $(\mathcal{P}_b(\mathbb{R}^d), d_{W_b})$ is a complete and separable metric space [2, Proposition 7.1.5]. We now suppose $b > 1$. While bounded subsets of $(\mathcal{P}_b(\mathbb{R}^d), d_{W_b})$ are not generally relatively compact in the b-Wasserstein metric [2, Remark 7.1.9], they are relatively compact with respect to d_{W_a} for $a < b$. Likewise, convergence in d_{W_b} can be characterized as follows (c.f [2, Remark 7.1.11]):

$$d_{W_b}(\mu_n, \mu) \to 0 \iff \mu_n \to \mu \text{ weak-}* \text{ in } \mathcal{P}(\mathbb{R}^d) \text{ and } \int |x|^b \, d\mu_n(x) \to \int |x|^b \, d\mu(x),$$

$$\iff \int f(x) \, d\mu_n(x) \to \int f(x) \, d\mu(x),$$

$$\text{for all } f \in C(\mathbb{R}^d) \text{ such that } |f(x)| \leqslant C(1 + |x - x_0|^b).$$

When $b = 2$, we abbreviate $d_W = d_{W_2}$.

2.2. Gradient Flows and their Γ-convergence.

We now briefly recall the notion of a *curve of maximal slope* in a compete metric space $(\mathcal{S}, d)$, which generalizes the concept of gradient flows outside the Riemannian context. We refer again to the book by Ambrosio, Gigli, and Savaré [2] for further details. A curve $u(t) : (a, b) \to \mathcal{S}$ is *2-absolutely continuous* if there exists $m \in L^2(a, b)$ so that

$$(2.3) \qquad\qquad d(u(t), u(s)) \leqslant \int_s^t m(r) \, dr \text{ for all } a < s \leqslant t < b.$$

We denote the space of 2-absolutely continuous curves by $AC^2([a, b], \mathcal{S})$.

For any 2-absolutely continuous curve, the limit

$$|u'(t)| = \lim_{s \to t} \frac{d(u(s), u(t))}{|s - t|}$$

exists for a.e. $t \in (a, b)$. Furthermore $m(t) := |u'(t)| \in L^2(a, b)$ satisfies (2.3) and for any $m \in L^2(a, b)$ satisfying (2.3), we have $|u'(t)| \leqslant m(t)$ for a.e. $t \in (a, b)$.

Given a functional $\mathsf{F} : \mathcal{S} \to (-\infty, +\infty]$ that is *proper*, i.e., $D(\mathsf{F}) = \{u \in \mathcal{S} : \mathsf{F}(u) < +\infty\} \neq \emptyset$, its upper gradient is a generalization of the modulus of the gradient from Euclidean space. Specifically, $g : \mathcal{S} \to [0, +\infty]$ is a *strong upper gradient* for F if for every $u \in AC^2([a, b], \mathcal{S})$ the function $g \circ u$ is measurable and

$$(2.4) \qquad\qquad |\mathsf{F}(u(t)) - \mathsf{F}(u(s))| \leqslant \int_s^t g(u(r))|u'|(r) \, dr \text{ for all } a < s \leqslant t < b.$$

When F is convex and lower semicontinuous, one example of a strong upper gradient is given by the metric local slope [2, Corollary 2.4.10] ,

$$(2.5) \qquad |\partial\mathsf{F}|(u) := \limsup_{v \to u} \frac{(\mathsf{F}(u) - \mathsf{F}(v))_+}{d(u,v)}.$$

Next, we recall the definition of a curve a maximal slope. A locally 2-absolutely continuous curve $u : (a,b) \to \mathcal{S}$ is a *curve of maximal slope* for F with respect to the strong upper gradient g if there exists a non-increasing function ϕ so that $\phi(t) = \mathsf{F} \circ u(t)$ for a.e. $t \in (a,b)$ and

$$(2.6) \qquad \phi'(t) \leqslant -\frac{1}{2}|u'|^2(t) - \frac{1}{2}g^2(u(t)) \text{ for a.e. } t \in (a,b).$$

Suppose $\mathsf{F} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\}$ is proper, lower semicontinuous, and $D(|\partial\mathsf{F}|) \subseteq \mathcal{P}_{2,ac}(\mathbb{R}^d)$. For any $\mu \in D(|\partial\mathsf{F}|)$, a map $\xi \in L^2(\mu)$ belongs to the *subdifferential* of F at μ if

$$(2.7) \qquad \mathsf{F}(\nu) - \mathsf{F}(\mu) \geqslant \int \langle \xi, \mathbf{t}_\mu^\nu - \mathrm{Id} \rangle d\mu + o(d_W(\mu,\nu)) \text{ for all } \nu \xrightarrow{d_W} \mu.$$

We denote this by $\xi \in \partial\mathsf{F}(\mu)$.

Remark 2.1 (subdifferential and metric slope). For any $\xi \in \partial\mathsf{F}(\mu)$, we have $\|\xi\|_{L^2(\mu)} \geqslant |\partial\mathsf{F}|(\mu)$.

When $\mathcal{S} = \mathcal{P}_2(\mathbb{R}^d)$ is endowed with the 2-Wasserstein metric d_W, a locally 2-absolutely continuous curve $\mu : (0,\infty) \to \mathcal{P}_2(\mathbb{R}^d)$ with $|\mu'| \in L^2_{\mathrm{loc}}(0,\infty)$ is called a *gradient flow* relative to the functional F if its velocity vector $v(t)$ satisfies

$$(2.8) \qquad -v(t) \in \partial\mathsf{F}(\mu(t)), \quad v(t) \in \mathrm{Tan}_{\mu(t)}\mathcal{P}_2(\mathbb{R}^d) \qquad \text{for a.e. } t \in (0,\infty).$$

The velocity vector field v is associated to μ through the continuity equation $\partial_t\mu + \nabla \cdot (v\mu) = 0$. If F is proper, lower semicontinuous, bounded below, and regular (see Definition 8.1) and its metric slope $|\partial\mathsf{F}|$ is a strong upper gradient, then $\mu(t)$ is a gradient flow of F if and only if $\mu(t)$ is a curve of maximal slope for $|\partial\mathsf{F}|$ [2, Theorem 11.1.3]. In particular, if $\mu(t)$ is a gradient flow of F, then $|\mu'|(t) = |\partial\mathsf{F}|(\mu(t))$ for almost every t.

With these definitions in hand, we now recall a general result of Serfaty on the Γ-convergence of gradient flows on a metric space. We state a mild variant of this result, similar to that used in [35, Theorem 5.6], which is a direct consequence of Serfaty's original proof.

Theorem 2.2 (cf. [64, Theorem 2]). *Let F_n and F be functionals defined on $(\mathcal{P}_2(\mathbb{R}^d), d_W)$ with strong upper gradients $|\partial\mathsf{F}_n|$ and $|\partial\mathsf{F}|$. Suppose that μ_n is a curve of maximal slope of F_n with well-prepared initial data $\mu_n(0)$, i.e., there exists $\mu(0) \in D(\mathsf{F})$ so that*

$$\mu_n(0) \xrightarrow{*} \mu(0) \quad and \quad \lim_{n \to \infty} \mathsf{F}_n(\mu_n(0)) = \mathsf{F}(\mu(0)).$$

If there exists some $\mu \in AC^2([0,T], \mathcal{P}_2(\mathbb{R}^d))$ so that $\mu_n(t) \xrightarrow{*} \mu(t)$ for all $t \in [0,T]$ and

    (i) $\liminf_{n \to \infty} \mathsf{F}_n(\mu_n(t)) \geqslant \mathsf{F}(\mu(t))$,
    (ii) $\liminf_{n \to \infty} \int_0^t |\mu_n'|^2(s)\,ds \geqslant \int_0^t |\mu'|^2(s)\,ds$,
    (iii) $\liminf_{n \to \infty} \int_0^t |\partial\mathsf{F}_n|^2(\mu_n(s))\,ds \geqslant \int_0^t |\partial\mathsf{F}|^2(\mu(s))\,ds$,

then μ is a curve of maximal slope of F and

$$\lim_{n \to \infty} \mathsf{F}_n(\mu_n(t)) = \mathsf{F}(\mu(t)) \quad for\ all\ \ t \in [0,T],$$

$$|\partial\mathsf{F}_n|(\mu_n(t)) \to |\partial\mathsf{F}|(\mu(t))\ and\ |\mu_n'|(t) \to |\mu'|(t)\ in\ L^2([0,T]).$$

For the 2-Wasserstein metric d_W the second criterion in Theorem 2.2 above holds independent of the choice of the energy functionals, and follows from the properties of the metric only, as the following elementary lemma shows. For lack of a reference, we include a proof.

Lemma 2.3 (Lower bound on metric derivatives). *Suppose μ_n and $\mu \in AC^2([0,T], \mathcal{P}_2(\mathbb{R}^d))$ for all $n \in \mathbb{N}$. If $\mu_n(t) \overset{*}{\rightharpoonup} \mu(t)$ in $\mathcal{P}_2(\mathbb{R}^d)$ for all $t \in [0,T]$, then*

$$\liminf_{n\to\infty} \int_0^s |\mu_n'|^2(t)\, dt \geqslant \int_0^s |\mu'|^2(t)\, dt. \text{ for all } s \in [0,T].$$

Proof. We may assume, without loss of generality, that there exists $0 \leqslant C < +\infty$ so that

$$C = \liminf_{n\to+\infty} \int_0^s |\mu_n'|^2(t)\, dt.$$

Choose a subsequence $|\mu_n'|(t)$ so that $\lim_{n\to+\infty} \int_0^s |\mu_n'|^2(t)\, dt = C$. Then $|\mu_n'|(t)$ is bounded in $L^2(0,s)$ as a sequence in $n \in \mathbb{N}$, so, up to a further subsequence, it is weakly convergent to some $\nu(t) \in L^2(0,s)$. Consequently, for any $0 \leqslant s_0 \leqslant s_1 \leqslant s$,

$$\lim_{n\to+\infty} \int_{s_0}^{s_1} |\mu_n'|(t)\, dt = \int_{s_0}^{s_1} \nu(t)\, dt.$$

By taking limits in the definition of the metric derivative and using the lower semicontinuity of d_W with respect to weak-* convergence,

$$d_W(\mu_n(s_0), \mu_n(s_1)) \leqslant \int_{s_0}^{s_1} |\mu_n'|(t)\, dt \quad \text{yields} \quad d_W(\mu(s_0), \mu(s_1)) \leqslant \int_{s_0}^{s_1} \nu(t)\, dt.$$

By [2, Theorem 1.1.2], this implies that $|\mu'|(t) \leqslant \nu(t)$ for a.e. $t \in (0,s)$. Thus, by the lower semicontinuity of the $L^2(0,s)$-norm with respect to weak convergence,

$$\liminf_{n\to+\infty} \int_0^s |\mu_n'|^2(t)\, dt \geqslant \int_0^s |\nu(t)|^2\, dt \geqslant \int_0^s |\mu'|^2(t)\, dt,$$

and we obtain the result. $\qquad\square$

3. Power-Law Interaction Potentials

In the present section, we prove that the interaction potentials described in the introduction satisfy our main hypotheses.

Theorem 3.1. *Suppose K is a power-law interaction potential of the form*

$$(3.1) \qquad K(x) = |x|^p/p \quad \text{or} \quad K(x) = |x|^q/q - |x|^p/p \quad \text{with} \quad 2 - d \leqslant p < q \leqslant 2,$$

where we adopt the convention $|x|^0/0 = \log(|x|)$. Then for all $m \geqslant m_0 > \max\{d/(d+p-1), 1\}$, K satisfies hypotheses (LSC), (GF1)–(GF3) for all $2 - d \leqslant p < q \leqslant 2$, and (ATT) for $2 - d \leqslant p < q \leqslant 2$ and $q > 0$.

Remark 3.2. Note that nonnegative combinations of the above potentials continue to satisfy the hypotheses, where the constraint on m depends on the most singular part of the potential.

We start by showing that power-law interaction potentials K satisfy hypothesis (LSC).

Proposition 3.3. *If $K(x) = |x|^p/p$ with $-d < p < 0$, then (LSC) holds for $r = -d/p$. If $K(x) = |x|^p/p$ with $p \geqslant 0$ or $K(x) = |x|^q/q - |x|^p/p$ with $-d < p < q$, then K satisfies (LSC) for all $r \in (1, +\infty)$.*

Proof. By definition, K is even and locally integrable. Suppose $K(x) = |x|^p/p$ for $p > 0$ or $K(x) = |x|^q/q - |x|^p/p$ with $-d < p < q$. Then K is lower semicontinuous and bounded from below; hence, (LSC) is satisfied with $K_a = K$ and $K_b = 0 \in L^r(\mathbb{R}^d)$ for all $r \in [1, +\infty)$.

Now, suppose $K(x) = |x|^0/0 = \log(|x|)$. Let $B = B_1(0)$ and define $K_a = K\chi_{\mathbb{R}^d \setminus B}$ and $K_b = K\chi_B$. Then K_a is continuous and bounded below and $K_b \in L^r(\mathbb{R}^d)$ for all $r \in [1, +\infty)$. Finally, suppose $K(x) = |x|^p/p$ with $-d \leqslant p < 0$. Then we have $K \in L^{-d/p,\infty}(\mathbb{R}^d)$, and (LSC) is satisfied with $K_a = 0$ and $K_b = K$. $\qquad\square$

We now show that power-law interaction potentials from Theorem 3.1 satisfy (ATT).

Proposition 3.4. *Suppose $K(x) = |x|^p/p$ with $-d < p \leqslant 2$ or $K(x) = |x|^q/q - |x|^p/p$ with $-d < p < q \leqslant 2$ and $q > 0$. Then K satisfies hypothesis (ATT).*

Proof. Suppose $K(x) = |x|^p/p$ with $p > 0$ or $K(x) = |x|^q/q - |x|^p/p$ with $-d < p < q \leqslant 2$ and $q > 0$. Then $\lim_{|x|\to\infty} K(x) = +\infty$, and (ATT)(ii) is satisfied. Similarly, for $K(x) = |x|^p/p$ with $p = 0$, K grows to infinity and satisfies hypothesis (ATT)(i). On the other hand, when $K(x) = |x|^p/p$ with $p < 0$, K is strictly increasing with $\lim_{|x|\to\infty} K(x) = 0$. Moreover, for $B = B_1(0)$, we have that $K \in L^a(\mathbb{R}^d \setminus B)$ for some fixed $1 < a < \infty$, $K \in C^1(\mathbb{R}^d \setminus \{0\})$, and $|\nabla K| \leqslant 1$ for $|x| > 1$. Therefore, K satisfies the hypothesis (ATT)(i). $\qquad\square$

Next we verify the hypotheses (GF1)–(GF3). As these are preserved under finite linear combinations of interaction potentials K, it suffices to show them for potentials of the form

$$(3.2) \qquad K(x) = |x|^p/p, \qquad 2 - d \leqslant p \leqslant 2.$$

We begin with a few results concerning potentials of this form and conclude with the proof of Theorem 3.1 at the end of the section.

Lemma 3.5. *Suppose $K(x) = |x|^p/p$ for $2 - d \leqslant p \leqslant 1$. Then for all $m_0 > d/(d + p - 1)$ and $\rho \in \mathcal{P}_2(\mathbb{R}^d) \cap L^{m_0}(\mathbb{R}^d)$ there exists $C_{d,p} > 0$ so that $\|\nabla K * \rho\|_\infty \leqslant C_{d,p}(1 + \|\rho\|_{m_0})$.*

Proof. Since $m_0 > d/(d + p - 1)$, its Hölder exponent satisfies $m_0' < d/(1 - p)$ and $\nabla K \in L^{m_0'}(B) \cap L^\infty(\mathbb{R}^d \setminus B)$, where $B = B_1(0) \subset \mathbb{R}^d$ is the unit ball. Thus,

$$\|\nabla K * \rho\|_\infty \leqslant \|(\nabla K\chi_B) * \rho\|_\infty + \|(\nabla K\chi_{\mathbb{R}^d \setminus B}) * \rho\|_\infty$$
$$\leqslant \|\nabla K\|_{L^{m_0'}(B)}\|\rho\|_{m_0} + \|\nabla K\|_{L^\infty(\mathbb{R}^d \setminus B)}\|\rho\|_1,$$

where χ_B denotes the characteristic function of B. $\qquad\square$

We now consider the hypothesis (GF1).

Proposition 3.6. *Suppose $K(x) = |x|^p/p$ for $2 - d \leqslant p \leqslant 2$. Then K satisfies hypothesis (GF1) for all $m_0 > \max\{d/(d + p - 1), 1\}$.*

Proof. When $p \leqslant 1$, we may use the uniform bound from Lemma 3.5 to conclude that for any $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, we have $\|\nabla K * \rho\|_{L^2(\nu)} \leqslant \|\nabla K * \rho\|_\infty \|\nu\|_1^{1/2} \leqslant C(\|\rho\|_{m_0} + 1) \leqslant C'(\|\rho\|_m + 1)$. Now, assume $1 < p \leqslant 2$. Then there exists $c > 0$ so that $|\nabla K(x)| \leqslant c(|x| + 1)$ for all $x \in \mathbb{R}^d$. Thus, by Minkowski's integral inequality and Jensen's inequality for the concave function $s \mapsto s^{1/2}$,

$$\|\nabla K * \rho\|_{L^2(\nu)} \leqslant \int \left(\int |\nabla K(x - y)|^2 \, d\rho(y) \right)^{1/2} d\nu(x)$$
$$\leqslant c \left(\iint (2|x|^2 + 2|y|^2 + 1) \, d\rho(y)d\nu(x) \right)^{1/2} \leqslant \sqrt{2}c(M_2(\rho)^{1/2} + M_2(\nu)^{1/2} + 1),$$

which gives the result. $\qquad\square$

We next turn to hypothesis (GF2).

Proposition 3.7. *Suppose $K(x) = |x|^p/p$ for $2 - d \leqslant p \leqslant 2$. Then K satisfies hypothesis (GF2) for all $m_0 > \max\{d/(d + p - 1), 1\}$.*

Proof. It suffices to estimate each component of the gradient $\nabla K * \rho = [\partial_i K * \rho]$ separately. Our approach is classical (cf. [60, Theorem 2.2]), extending known results about continuity properties of singular integrals to interaction potentials with at most quadratic growth at infinity.

Let $R = |x - y|$. Then,

$$|\partial_i K * \rho(x) - \partial_i K * \rho(y)| = \left| \int \left(\partial_i K(x - z) - \partial_i K(y - z) \right) d\rho(z) \right|$$

$$= \left| \int_{B_{2R}(x)} + \int_{\mathbb{R}^d \setminus B_{2R}(x)} \right| \leqslant \left| \int_{B_{2R}(x)} \right| + \left| \int_{\mathbb{R}^d \setminus B_{2R}(x)} \right| =: \mathrm{I} + \mathrm{II}$$

We begin by estimating I. If $p = 2$, let $\beta = 1$. Otherwise, choose $\beta \in (0, 1]$ so that $m_0 > d/(d + p - 1 - \beta) > 1$. Define $r := d/(d + p - 1 - \beta) \in [1, m_0)$, and let r' be the conjugate index of r. Since $p - 1 + (d/r') = \beta \in (0, 1]$,

$$\mathrm{I} \leqslant \int_{B_{2R}(x)} |\partial_i K(x - z)| \, d\rho(z) + \int_{B_{2R}(x)} |\partial_i K(y - z)| \, d\rho(z)$$

$$\leqslant \int_{B_{2R}(x)} |x - z|^{p-1} \, d\rho(z) + \int_{B_{3R}(y)} |y - z|^{p-1} \, d\rho(z)$$

$$\leqslant \|\rho\|_r \left\| |x - \cdot|^{p-1} \right\|_{L^{r'}(B_{2R}(x))} + \|\rho\|_r \left\| |y - \cdot|^{p-1} \right\|_{L^{r'}(B_{3R}(y))}$$

$$\leqslant C\|\rho\|_r |x - y|^\beta,$$

for $C = C_{d,p,r}$. Now, we estimate II. For $x_\alpha := \alpha x + (1 - \alpha)y$, $\alpha \in [0, 1]$, we have

$$\mathrm{II} = \left| \int_{\mathbb{R}^d \setminus B_{2R}(x)} \int_0^1 \frac{d}{d\alpha} \partial_i K(x_\alpha - z) \, d\alpha d\rho(z) \right|$$

$$= \left| \int_{\mathbb{R}^d \setminus B_{2R}(x)} \int_0^1 \langle \nabla \partial_i K(x_\alpha - z), y - x \rangle \, d\alpha d\rho(z) \right|$$

$$\leqslant C |x - y| \max_{\alpha \in [0,1]} \int_{\mathbb{R}^d \setminus B_{2R}(x)} |x_\alpha - z|^{p-2} \, d\rho(z).$$

If $p = 2$, the integral is bounded by 1 and $\beta = r = 1$. Thus, $\mathrm{II} \leqslant C|x - y|^{p-1+(d/r')} = C|x - y|^\beta$. If $p < 2$, then $p - 2 + (d/r') \leqslant 0$, and we may bound the integral as follows,

$$\int_{\mathbb{R}^d \setminus B_{2R}(x)} |x_\alpha - z|^{p-2} \, d\rho(z) \leqslant \|\rho\|_r \left(\int_{\mathbb{R}^d \setminus B_{2R}(x)} |x_\alpha - z|^{(p-2)r'} \, dz \right)^{1/r'}$$

$$\leqslant \|\rho\|_r \left(\int_{\mathbb{R}^d \setminus B_R(x_\alpha)} |x_\alpha - z|^{(p-2)r'} \, dz \right)^{1/r'}$$

$$\leqslant C\|\rho\|_r |x - y|^{p-2+(d/r')},$$

Therefore,

$$\mathrm{I} + \mathrm{II} \leqslant C\|\rho\|_r |x - y|^{p-1+(d/r')} \leqslant C(\|\rho\|_m + 1)\psi\big(|x - y|\big),$$

for $\psi(s) = s^{p-1+(d/r')} = s^\beta$, $\beta \in (0, 1]$, which completes the proof of (GF2). $\qquad \square$

In order to show that power-law interaction potentials satisfy property (GF3), we begin with the following estimate, quantifying the stability of $\nabla K * \rho$ in the 1-Wasserstein metric when K is sufficiently regular.

Lemma 3.8. *Suppose $K(x) = |x|^p/p$ for $1 < p \leqslant 2$. Then there exists $C > 0$ so that, for all $\rho, \nu \in \mathcal{P}_1(\mathbb{R}^d)$,*

$$\|\nabla K * \rho - \nabla K * \nu\|_\infty \leqslant C(d_{W_1}(\rho, \nu))^{(p-1)}.$$

Proof. Let $\gamma_0 \in \mathcal{C}_0(\rho, \nu)$ be the optimal transport plan from ρ to ν. Using $|\nabla K(v) - \nabla K(w)| \leqslant C|v - w|^{p-1}$ and the concavity of the right hand side, we obtain

$$\|\nabla K * \rho - \nabla K * \nu\|_\infty = \sup_{x \in \mathbb{R}^d} \left| \iint (\nabla K(x-y) - \nabla K(x-z)) \, d\gamma_0(y, z) \right|$$

$$\leqslant C \iint |z - y|^{p-1} \, d\gamma_0(y, z) \leqslant C \left(\int |z - y| \, d\gamma_0(y, z) \right)^{(p-1)},$$

which completes the proof. $\qquad \square$

We now use the previous estimate to quantify the stability of $\nabla K * \rho$ in the b-Wasserstein metric for all $b \in [1, 2]$ and for general attractive power-law potentials. This generalizes a result of Loeper [54, Theorem 4.4] to general power-law potentials, L^p spaces, and b-Wasserstein metrics for $b \leqslant 2$. This generalization plays a key role in our proof of the Γ-convergence of gradient flows, since the 2-Wasserstein gradient flow structure merely provides compactness in b-Wasserstein metrics for $b < 2$. To obtain convergence of the subdifferentials of the interaction energies, we require continuity of $\nabla K * \rho$ with respect to weaker Wasserstein metrics.

Proposition 3.9. *Suppose $K(x) = |x|^p/p$ for $2 - d \leqslant p \leqslant 2$ and fix β so that*

$$(p-1)_+ < \beta \leqslant 1 \text{ for } p < 2 \quad or \quad \beta = 1 \text{ for } p = 2.$$

Then for all $0 \leqslant \epsilon < 1$, there exists $C > 0$, depending on d, p, ϵ, and β, so that

$$\|\nabla K * \rho - \nabla K * \nu\|_{p_*} \leqslant C \max\left\{ \|\rho\|_{m_*}, \|\nu\|_{m_*} \right\}^{(1-\epsilon)/(2-\epsilon)} d_{W_{2-\epsilon}}(\rho, \nu)$$

for all $\rho, \nu \in \mathcal{P}_2(\mathbb{R}^d) \cap L^{m_*}(\mathbb{R}^d)$, where

$$m_* = \frac{d}{d + p - 1 - \beta} \quad and \quad p_* = \frac{d(2 - \epsilon)}{3 - p - \beta - \epsilon(1 - \beta)}.$$

Proof. Let $b = 2 - \epsilon \in (1, 2]$, and let $\mathbf{t}_\rho^\nu$ be an optimal transport map from ρ to ν with respect to the b-Wasserstein metric. If $p = 2$, then $m_* = 1$, $p_* = +\infty$, and by Lemma 3.8,

$$\|\nabla K * \rho - \nabla K * \nu\|_{p_*} \leqslant C d_{W_1}(\rho, \nu) \leqslant C d_{W_b}(\rho, \nu).$$

Now, suppose $p < 2$, so $p_* \in (1, +\infty)$. Let $\rho_\alpha := ((1 - \alpha)\mathbf{t}_\rho^\nu + \alpha \, \mathrm{Id})_{\#}\rho$, $\alpha \in [0, 1]$, be the constant speed geodesic in the b-Wasserstein metric, and let $\mathbf{t}_{\rho_\alpha}^{\rho_1}$ and $\mathbf{t}_{\rho_\alpha}^{\rho_0}$ the optimal transport maps from intermediate points along the geodesics to the endpoints, $\alpha \in (0, 1)$ [2, Lemma

7.2.1]. Then, by Minkowski's integral inequality,

$$\|\nabla K * \rho - \nabla K * \nu\|_{p_*} = \left\|\int_0^1 \frac{d}{d\alpha} \nabla K * \rho_\alpha \, d\alpha\right\|_{p_*} \leqslant \int_0^1 \left\|\frac{d}{d\alpha} \nabla K * \rho_\alpha\right\|_{p_*} d\alpha$$

$$= \int_0^1 \left\|\int D^2 K(\cdot - (1-\alpha)y - \alpha \mathbf{t}_\rho^\nu(y))(\mathbf{t}_\rho^\nu(y) - y)\, d\rho_0(y)\right\|_{p_*} d\alpha$$

$$= \int_0^1 \left\|\int D^2 K(\cdot - y)(\mathbf{t}_{\rho_\alpha}^{\rho_1}(y) - \mathbf{t}_{\rho_\alpha}^{\rho_0}(y))\, d\rho_\alpha(y)\right\|_{p_*} d\alpha = \int_0^1 \left\|D^2 K * \left[(\mathbf{t}_{\rho_\alpha}^{\rho_1} - \mathbf{t}_{\rho_\alpha}^{\rho_0})\rho_\alpha\right]\right\|_{p_*} d\alpha.$$

By Sobolev's inequality for Riesz potentials ($p > 2 - d$; cf. [60, Section 4.2]) and the Calderón Zygmund inequality ($p = 2 - d$; cf. [66, Theorem V.1]) and Hölder's inequality,

$$\|D^2 K * [(\mathbf{t}_{\rho_\alpha}^{\rho_1} - \mathbf{t}_{\rho_\alpha}^{\rho_0})\rho_\alpha]\|_{p_*} \leqslant C_{d,p,p_*} \|(\mathbf{t}_{\rho_\alpha}^{\rho_1} - \mathbf{t}_{\rho_\alpha}^{\rho_0})\rho_\alpha\|_l$$

$$\leqslant C_{d,p,p_*} \left\|(\mathbf{t}_{\rho_\alpha}^{\rho_1} - \mathbf{t}_{\rho_\alpha}^{\rho_0})\rho_\alpha^{1/b}\right\|_b \|\rho_\alpha^{(b-1)/b}\|_r$$

$$= C_{d,p,p_*} d_{W_b}(\rho, \nu) \|\rho_\alpha\|_{m_*}^{(b-1)/b}$$

for

$$l = \frac{(2-\epsilon)d}{2d + p - 1 - \beta + \epsilon(\beta - d - p + 1)} \in (1, 2] \quad \text{and} \quad r = \frac{d(2-\epsilon)}{(d + p - 1 - \beta)(1-\epsilon)} \in [2, +\infty].$$

Finally, by convexity of L^p-norms along b-Wasserstein geodesics [2, Proposition 9.3.9],

$$\|\rho_\alpha\|_{m_*} \leqslant (1-\alpha)\|\rho\|_{m_*} + \alpha\|\nu\|_{m_*} \leqslant \max\{\|\rho\|_{m_*}, \|\nu\|_{m_*}\},$$

which, combined with the previous inequalities gives the result.   □

Finally, we use the result of Proposition 3.9 to show that power-law interaction potentials satisfy hypothesis (GF3).

Proposition 3.10. *Suppose $K(x) = |x|^p/p$ for $2 - d \leqslant p \leqslant 2$. Then K satisfies hypothesis (GF3) for all $m_0 > \max\{d/(d + p - 1), 1\}$. In particular, for all $0 \leqslant \epsilon < 1$, there exists $C_\epsilon = C(p, d, m_0, \epsilon) > 0$ and $\alpha_\epsilon = \alpha(p, d, m_0, \epsilon) \in (0, 1]$ so that*

$$\|\nabla K * \rho - \nabla K * \nu\|_{L^2(\mu)} \leqslant C_\epsilon \Big(1 + \|\rho\|_{m_0} + \|\nu\|_{m_0} + \|\mu\|_{m_0}\Big) d_{W_{2-\epsilon}}^{\alpha_\epsilon}(\rho, \nu).$$

Remark 3.11 (range of m_0 and ϵ). For m_0 sufficiently large and $\epsilon = 0$, we obtain $\alpha = 1$, and this proposition reduces to [42, Assumption 4.1(v)]. Consequently, the main contribution of this new estimate is that it lowers the range of admissible values of m_0 and strengthens the Wasserstein metric in the estimate. This extension is crucial is the case $p = 2 - d$, $d > 2$, since previous works required $m_0 = +\infty$, while this new estimate allows all $m_0 > d$. This extension is also crucial in that it provides stability with respect to weaker Wasserstein metrics when $\epsilon > 0$. Such a stability result is needed in our proof of Γ-convergence of the gradient flows, where we merely obtain compactness of the gradient flows in b-Wasserstein metrics for $1 \leqslant b < 2$.

Proof. When $1 \leqslant p \leqslant 2$, Lemma 3.8 and the fact that $\mu \in \mathcal{P}(\mathbb{R}^d)$ ensures

$$\|\nabla K * \rho - \nabla K * \nu\|_{L^2(\mu)} \leqslant \|\nabla K * (\rho - \nu)\|_\infty \leqslant d_{W_1}(\rho, \nu)^{p-1} \leqslant d_{W_{2-\epsilon}}(\rho, \nu)^{p-1}$$

which gives the result.

Now, suppose that $p < 1$. Let

$$\beta = \min\left\{d + p - 1 - d/m_0, 1\right\} \in ((p-1)_+, 1].$$

Note that $\beta = 1$ if and only if $m_0 \geqslant \frac{d}{d+p-2}$. Fix $0 \leqslant \epsilon < 1$ and define m_* and p_* as in Theorem 3.9, so that $m_0 \geqslant m_* > 1$. Since $2m_*/(m_*-1) \geqslant p_* \geqslant 1$, there exists $\alpha_\epsilon = \alpha(p,d,m_0,\epsilon) \in (0,1]$ so that

$$\alpha \frac{2m_*}{m_*-1} = p_*.$$

By Hölder's inequality and interpolation of L^p-norms,

$$\|\nabla K * \rho - \nabla K * \nu\|_{L^2(\mu)} \leqslant \|\nabla K * (\rho - \nu)\|_{2m_*/(m_*-1)} \|\mu\|_{m_*}^{1/2}$$
$$\leqslant \|\nabla K * (\rho - \nu)\|_{p_*}^{\alpha} \|\nabla K * (\rho - \nu)\|_{\infty}^{1-\alpha} \|\mu\|_{m_*}^{1/2}$$

Lemma 3.5 ensures $\|\nabla K * (\rho - \nu)\|_\infty \leqslant C_{d,p}(1 + \|\rho\|_{m_*} + \|\nu\|_{m_*})$. Therefore, applying this and Proposition 3.9 gives

$$\|\nabla K * \rho - \nabla K * \nu\|_{L^2(\mu)}$$
$$\leqslant C \|\nabla K * (\rho - \nu)\|_{p_*}^{\alpha} \left(1 + \|\rho\|_{m_*} + \|\nu\|_{m_*}\right)^{1-\alpha} \|\mu\|_{m_*}^{1/2}$$
$$\leqslant C d_{W_{2-\epsilon}}(\rho,\nu)^{\alpha} \max\{\|\rho\|_{m_*}, \|\nu\|_{m_*}\}^{\frac{\alpha(1-\epsilon)}{2-\epsilon}} \left(1 + \|\rho\|_{m_*} + \|\nu\|_{m_*}\right)^{1-\alpha} \|\mu\|_{m_*}^{1/2}.$$

Simplifying and using that $m_0 \geqslant m_*$, we conclude the result for $\psi(s) = s^\alpha$. □

We conclude this section with the proof of Theorem 3.1.

Proof of Theorem 3.1. Hypothesis (LSC) follows from Proposition 3.3. Hypothesis (GF1)–(GF3) are preserved under finite linear combinations of interaction potentials K, so it suffices to show them for power-law potentials of the form $K(x) = |x|^p/p$ for $m_0 > \max\{d/(d+p-1),1\}$. Hypothesis (GF1) follows from Proposition 3.6. To see hypotheses (GF2)–(GF3), let $\psi_1(s) = s^\beta$ be the modulus of continuity from Proposition 3.7 and let $\psi_2(s) = s^\alpha$ be the modulus of continuity from Proposition 3.10. Then K satisfies hypotheses (GF2)–(GF3) with $\psi(s) := \psi_1(s) + \psi_2(s)$. □

Remark 3.12. (restriction to power-law potentials with $p \geqslant 2 - d$) As shown in Propositions 3.3 and 3.4, we merely require $p > -d$ for the corresponding energies E_m to satisfy hypotheses (LSC) and (ATT). (Furthermore, by Proposition 4.1, hypothesis (LSC) is sufficient to ensure that E_m is lower semicontinuous and bounded below; hence its gradient flows exist.) Still, in order to characterize the subdifferentials of the gradient flows of E_m and study their limits as $m \to +\infty$, we need control over derivatives of $K * \rho$ when $\rho \in L^m(\mathbb{R}^d)$. Consequently, our results on Γ-convergence of the gradient flows require that our energies E_m satisfy hypotheses (GF1)–(GF3), which hold merely for $p \geqslant 2 - d$.

4. Energies, Gradient Flows, and Aggregation-Diffusion Equations

In this section, we develop several fundamental properties of the energies and E_m and E_∞, prove existence and (in some cases) uniqueness of their gradient flows. We also rigorously connect these gradient flows to aggregation-diffusion equations. In the process, we extend the well-posedness theory for such equations and, in some cases, obtain sharper estimates on solutions than has been previously obtained by pure PDE methods; see Remarks 1.8 and 1.9.

The key obstacle in analysis of the energies E_m and E_∞ is to quantify the competing effects of the interaction, diffusion, and height constraint terms. We denote the diffusion and height constraint parts of the energies by

$$(4.1) \quad \mathsf{S}_m(\rho) := \begin{cases} \dfrac{1}{m-1} \displaystyle\int \rho^m \, dx & \text{if } \rho \in L^1(\mathbb{R}^d) \cap L^m(\mathbb{R}^d), \\ +\infty & \text{otherwise,} \end{cases} \quad \mathsf{S}_\infty(\rho) := \begin{cases} 0 & \text{if } \|\rho\|_\infty \leqslant 1, \\ +\infty & \text{otherwise,} \end{cases}$$

and the interaction part by

$$\mathsf{K}(\rho) := \frac{1}{2} \iint K(x-y)\, d\rho(x) d\rho(y) \tag{4.2}$$

so that $\mathsf{E}_m = \mathsf{S}_m + \mathsf{K}$ and $\mathsf{E}_\infty = \mathsf{S}_\infty + \mathsf{K}$.

4.1. **Basic Properties of E_m and E_∞.** We now develop some basic properties of the energy functionals E_m and E_∞. (For ease of notation, we often consider both energies at the same time by proving properties for E_m, allowing $m = +\infty$.) The results in this section merely rely on hypothesis (LSC). In the case of attractive power-law interaction potentials, this is equivalent to requiring that we are in the diffusion dominated regime. (See Remark 1.1.)

We first show that the energies are bounded below and that an upper bound on $\mathsf{E}_m(\rho)$ implies an upper bound on $\|\rho\|_m$.

Proposition 4.1. *Suppose K satisfies hypothesis (LSC) and $m \in [m_0, +\infty]$. Then E_m is bounded below, uniformly in m, and there exists $C_r > 0$ s.t.*

$$\|\rho\|_m^{1+\frac{1}{r}} \leqslant \mathsf{E}_m(\rho) + C_r. \tag{4.3}$$

Proof. First, we show there exists $C_r > 0$ so that, for all $m \in [m_0, +\infty]$ and $\rho \in L^m(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^d)$,

$$\frac{1}{2} \int (K * \rho)\, d\rho \geqslant -C_r(1 + \|\rho\|_m^{1+\frac{1}{r}}). \tag{4.4}$$

By hypothesis (LSC) and the weak Young inequality (cf. [51, Section 4.3], [10, Lemma 4])

$$\frac{1}{2} \int (K * \rho)\, d\rho = \frac{1}{2} \int (K_a * \rho)\, d\rho + \frac{1}{2} \int (K_b * \rho)\, d\rho \geqslant -\frac{1}{2}\|K_a\|_\infty - \|K_b\|_{L^{r,\infty}} \|\rho\|_{1+\frac{1}{r}}^{1+\frac{1}{r}}.$$

Since $m \geqslant m_0 \geqslant 1 + \frac{1}{r}$ and $\rho \in \mathcal{P}(\mathbb{R}^d)$, interpolating $L^p(\mathbb{R}^d)$ norms,

$$\|\rho\|_{1+\frac{1}{r}}^{1+\frac{1}{r}} \leqslant \|\rho\|_m^{\frac{m}{r(m-1)}} \leqslant (1 + \|\rho\|_m)^{\frac{m}{r(m-1)}} \leqslant (1 + \|\rho\|_m)^{1+\frac{1}{r}} \leqslant C_r \left(1 + \|\rho\|_m^{1+\frac{1}{r}} \right).$$

Combining the two previous inequalities shows (4.4).

We now show inequality (4.3). For $m = +\infty$, this follows from inequality (4.4) and the definition of E_∞. Suppose $m \in [m_0, +\infty)$. By definition of E_m and inequality (4.4), for all $\rho \in D(\mathsf{E}_m)$,

$$\mathsf{E}_m(\rho) \geqslant \frac{1}{m-1}\|\rho\|_m^m - C(1 + \|\rho\|_m^{1+\frac{1}{r}}) = \|\rho\|_m^{1+\frac{1}{r}} \left(\frac{1}{m-1}\|\rho\|_m^{m-1-\frac{1}{r}} - C \right) - C. \tag{4.5}$$

First, assume

$$\frac{1}{m-1}\|\rho\|_m^{m-1-\frac{1}{r}} - C \geqslant 1. \tag{4.6}$$

Then, $\|\rho\|_m^{1+\frac{1}{r}} \leqslant \mathsf{E}_m(\rho) + C$ and (4.3) holds. Alternatively, suppose (4.6) does not hold. Then,

$$\|\rho\|_m \leqslant ((1+C)(m-1))^{1/(m-1-\frac{1}{r})}.$$

Since $m \geqslant m_0 > 1 + \frac{1}{r}$, there exists $C_1 = C_1(C, r) > 0$ so that $\|\rho\|_m^{1+\frac{1}{r}} \leqslant C_1$.

To conclude (4.3), it suffices to show that E_m is bounded below, uniformly in $m \geqslant m_0$. We will show it is bounded below on the set $\{\rho \in D(\mathsf{E}_m) : \mathsf{E}_m(\rho) \leqslant 1\}$. By the previous inequalities, on this set, we have $\|\rho_m\|_m^{1+\frac{1}{r}} \leqslant 1 + C + C_1$. Combining this with inequality (4.5), we obtain that the energy is uniformly bounded below. $\square$

We now turn to further properties of the energies. For all $m \in [1, +\infty]$, the energies S_m are proper, lower semicontinuous with respect to weak-* convergence, and convex in the 2-Wasserstein metric (cf. [2, Proposition 9.3.9], [42, Proposition 4.5]). Likewise, for all $p \in [1, +\infty)$ the $L^p(\mathbb{R}^d)$ norms

$$(4.7) \qquad \|\rho\|_p := \begin{cases} \|\rho\|_p & \text{if } \rho \in L^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d), \\ +\infty & \text{otherwise,} \end{cases}$$

are also proper, lower semicontinuous with respect to weak-* convergence, and convex in the 2-Wasserstein metric.

We now show that the interaction energy is also lower semicontinuous with respect to weak-* convergence, on $L^{m_0}(\mathbb{R}^d)$-bounded sets.

Proposition 4.2 (lower semicontinuity of interaction energy). *Suppose K satisfies hypothesis (LSC). If $\rho_n \overset{*}{\rightharpoonup} \rho$ and $\sup_{n \in \mathbb{N}} \|\rho_n\|_{m_0} < +\infty$, then $\liminf_{n \to +\infty} \mathsf{K}(\rho_n) \geqslant \mathsf{K}(\rho)$.*

Proof. First, note that since $\rho_n \overset{*}{\rightharpoonup} \rho$, we have also have $\rho_n \otimes \rho_n \overset{*}{\rightharpoonup} \rho \otimes \rho$. Since K_a is lower semicontinuous and bounded below, by Portmanteau theorem (see [68, Theorem 1.3.4]),

$$\liminf_{n \to +\infty} \frac{1}{2} \iint K_a(x - y) \, d\rho_n(x) d\rho_n(y) \geqslant \frac{1}{2} \iint K_a(x - y) \, d\rho(x) d\rho(y).$$

Now, we consider K_b. For any $k > 0$, define $K_b \wedge (-k) = \min\{K_b, -k\}$. Since $K_b : \mathbb{R}^d \to [-\infty, +\infty]$ is lower-semicontinuous, $K_b \wedge (-k)$ is lower semicontinuous and bounded below for all $k > 0$. Furthermore, since $K_b \in L^{r,\infty}(\mathbb{R}^d)$, if we define $S_k := \{x : K_b(x) \leqslant -k\}$, by the weak Young inequality (cf. [51, Section 4.3], [10, Lemma 4]), for any $m_* \in (1 + \frac{1}{r}, 2)$ with $m_* \leqslant m_0$,

$$\left| \int ((K_b \wedge (-k)) * \rho_n) \, d\rho_n - \int (K_b * \rho_n) \, d\rho_n \right| \leqslant \|K_b \chi_{S_k}\|_{L^{r_*,\infty}(\mathbb{R}^d)} \|\rho_n\|_{m_*}^{m_*}, \quad r_* = (m_* - 1)^{-1}$$

Since we assume $\sup_{n \in \mathbb{N}} \|\rho_n\|_{m_0} < +\infty$ and $m_* \leqslant m_0$, the second term is bounded uniformly in $n \in \mathbb{N}$. Likewise, since $r_* < r$, by definition of $L^{r,\infty}(\mathbb{R}^d)$,

$$\|K_b \chi_{S_k}\|_{L^{r_*,\infty}(\mathbb{R}^d)} = \sup_{\lambda > 0} \lambda |\{|K_b \chi_{S_k}| > \lambda\}|^{1/r_*} \leqslant \sup_{\lambda \geqslant k} \lambda |\{|K_b| > \lambda\}|^{1/r_*}$$

$$= \sup_{\lambda \geqslant k} \lambda^{(r_* - r)/r_*} \left(\lambda |\{|K_b| > \lambda\}|^{1/r} \right)^{r/r_*} \leqslant k^{(r_* - r)/r_*} \|K_b\|_{L^{r,\infty}(\mathbb{R}^d)}^{r/r_*} \xrightarrow{k \to +\infty} 0.$$

Therefore, for all $\epsilon > 0$, we may choose $k > 0$ so that

$$\liminf_{n \to +\infty} \frac{1}{2} \int (K_b * \rho_n) \, d\rho_n \geqslant \liminf_{n \to +\infty} \frac{1}{2} \iint (K_b(x - y) \wedge (-k)) \, d\rho_n(x) d\rho_n(y) - \epsilon$$

$$\geqslant \frac{1}{2} \int ((K_b \wedge (-k)) * \rho) \, d\rho - \epsilon \geqslant \frac{1}{2} \int (K_b * \rho) \, d\rho - \epsilon.$$

Since $\epsilon > 0$ was arbitrary and $K = K_a + K_b$, we obtain the result. $\qquad \square$

We conclude this section by applying the previous proposition to show that E_m is proper, lower semicontinuous, and bounded below.

Proposition 4.3. *Suppose K satisfies hypothesis (LSC) and $m \in [m_0, +\infty]$. Then E_m is proper, lower semicontinuous with respect to weak-* convergence of probability measures, and bounded below.*

Proof. The fact that E_m is bounded below is an immediate consequence of Proposition 4.1 and the fact that $\mathsf{S}_m \geqslant 0$. We now show E_m is proper. By hypothesis (LSC), $K \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Therefore, evaluating the energy E_m on there characteristic function of a ball B of volume 1, we have

$$\mathsf{E}_m(\chi_B) \leqslant \frac{1}{m_0 - 1} + \frac{1}{2} \iint_{B \times B} K(x - y)\, dxdy < +\infty.$$

Hence, E_m is proper.

We conclude by proving that E_m is lower semicontinuous. Suppose that $\rho_n \overset{*}{\rightharpoonup} \rho$. Without loss of generality, we may assume that $\liminf_{n \to +\infty} \mathsf{E}_m(\rho_n) < +\infty$. Taking a subsequence ρ_n so that the $\liminf$ is attained, we may also assume that $\sup_{n \in \mathbb{N}} \mathsf{E}_m(\rho_n) < +\infty$. Applying Proposition 4.1 and interpolation of L^p norms for $1 < m_0 \leqslant m$, we obtain that $\sup_{n \in \mathbb{N}} \|\rho_n\|_{m_0} < +\infty$. Therefore, by Proposition 4.2, $\liminf_{n \to +\infty} \mathsf{K}(\rho_n) \geqslant \mathsf{K}(\rho)$. Since [58, Corollary 3.5] ensures $\liminf_{n \to +\infty} \mathsf{S}_m(\rho_n) \geqslant \mathsf{S}_m(\rho)$, this gives the result. $\qquad\square$

Remark 4.4 (sharpness of condition $m \geqslant m_0$). The condition $m \geqslant m_0$ in Proposition 4.3 is sharp to ensure that the energy E_m is lower semicontinuous with respect to weak-* convergence. In particular, for all $\epsilon > 0$, there exists $K(x)$ and m_0 satisfying (LSC) and $m \in (m_0 - \epsilon, m_0)$ so that the energy E_m is not lower semicontinuous. For example, we may take $K(x) = |x|^p/p$ for $-d < p < 0$, $m_0 = 1 - (p/d) + \epsilon/4$, and $m = 1 - (p/d) - \epsilon/4$. (We assume, without loss of generality, that $\epsilon > 0$ is sufficiently small so that $m > 1$.) By Proposition 3.3, K and m_0 satisfy hypothesis (LSC). For any $\rho_* \in \mathcal{P}_2(\mathbb{R}^d)$ we may consider its sequence of dilations $\rho_\lambda = \lambda^d \rho_*(\lambda x)$, which converges in the weak-* topology as $\lambda \to +\infty$ to a Dirac mass at the origin δ_0. Along this sequence the energies E_m satisfy $\lim_{\lambda \to +\infty} \mathsf{E}_m(\rho_\lambda) = -\infty$ and $\mathsf{E}_m(\delta_0) = +\infty$ (see, e.g. [30, equations (12)-(13)]). Therefore, the energy E_m is not lower semicontinuous in the weak-* topology.

4.2. **Subdifferentials and Gradient Flows.** We now characterize the minimal elements of the subdifferential of E_m, $m \in [m_0, +\infty)$, and identify elements belonging to the subdifferential of E_∞. We defer our full characterization of minimal elements of the subdifferential of E_∞ to the proof of Theorem 1.10 in section 6. Following these results on the subdifferentials, we prove that gradient flows of E_m and E_∞ exist and provide conditions under which they are unique. Throughout this section, we use hypotheses (LSC), (GF1)–(GF3) on the interaction potential and suppose $m \geqslant m_0$.

In order to analyze the subdifferentials of E_m and E_∞, we begin with the following lemma, which bounds the variation of the nonlocal interaction energy K along measures in $L^m(\mathbb{R}^d)$. This lemma extends [42, Proposition 4.6], where (GF1), (GF2), and (GF3) generalize [42, Assumption 4.1]. We defer its proof to appendix section 8.2.

Proposition 4.5. *Suppose K satisfies (GF1)–(GF3) and $\rho_0, \rho_1 \in L^m(\mathbb{R}^d)$ for $m \geqslant m_0$. Then*

$$\left| \mathsf{K}(\rho_1) - \mathsf{K}(\rho_0) - \int \left\langle \nabla K * \rho_0, \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right\rangle d\rho_0 \right| \leqslant f(\rho_0, \rho_1)\, \psi(d_W(\rho_0, \rho_1))\, d_W(\rho_0, \rho_1)$$

for $f(\rho_0, \rho_1) = C'(1 + \|\rho_0\|_m + \|\rho_1\|_m)$ and $C' = C'(d_W(\rho_0, \rho_1)) > 0$ is an increasing function of the distance from ρ_0 to ρ_1.

We now apply the previous proposition to obtain the following generalization of Ambrosio, Gigli, and Savaré's characterization of the subdifferential for λ-convex energies to the energies E_m and E_∞.

Proposition 4.6. *Suppose K satisfies (LSC), (GF1)–(GF3), $m \in [m_0, +\infty]$, and $\rho \in D(\mathsf{E}_m)$. Then $\xi \in L^2(\rho)$ belongs to $\partial \mathsf{E}_m(\rho)$ if and only if*

$$(4.8) \qquad \mathsf{E}_m(\nu) - \mathsf{E}_m(\rho) \geqslant \int \langle \xi, \mathbf{t}_\rho^\nu - \mathrm{Id} \rangle \, d\rho - f(\rho, \nu)\psi(d_W(\rho,\nu))d_W(\rho,\nu), \ \forall \nu \in D(\mathsf{E}_m),$$

for $f(\rho, \nu)$ as in Proposition 4.5.

Proof. First, suppose $\xi \in L^2(\rho)$ satisfies inequality (4.8). We show that it satisfies the subdifferential inequality (2.7). For any sequence $\nu_n \to \rho$, $\nu_n \in D(\mathsf{E}_m)$, we may assume without loss of generality that $\sup_n \mathsf{E}_m(\nu_n) < +\infty$, or else the subdifferential inequality (2.7) is satisfied trivially. Hence, by Proposition 4.1, $\sup_n \|\nu_n\|_m < +\infty$. Therefore, $f(\rho, \nu_n)$ is uniformly bounded as $\nu_n \to \rho$. Consequently, by the definition of the subdifferential (2.7), $\xi \in \partial \mathsf{E}_m(\rho)$.

Now, suppose $\xi \in L^2(\rho)$ belongs to $\partial \mathsf{E}_m(\rho)$ and $\nu \in D(\mathsf{E}_m)$. We show that inequality (4.8) holds. Let $\rho_\alpha = ((1 - \alpha)\,\mathrm{Id} + \alpha \mathbf{t}_\rho^\nu)\#\rho$ be the Wasserstein geodesic from ρ to ν. Then by definition of the subdifferential, inequality (2.7),

$$\frac{d}{d\alpha}\mathsf{E}_m(\rho_\alpha)\bigg|_{\alpha=0} = \lim_{\alpha \to 0} \frac{\mathsf{E}_m(\rho_\alpha) - \mathsf{E}_m(\rho)}{\alpha} \geqslant \int \langle \xi, \mathbf{t}_\rho^\nu - \mathrm{Id} \rangle \, d\rho.$$

By Proposition 4.5,

$$\frac{d}{d\alpha}\mathsf{K}(\rho_\alpha)\bigg|_{\alpha=0} = \int \left\langle \nabla K * \rho_0, \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right\rangle d\rho_0 \leqslant \mathsf{K}(\nu) - \mathsf{K}(\rho) + f(\rho, \nu)\,\psi(d_W(\rho,\nu))\,d_W(\rho,\nu).$$

Likewise, by the convexity of S_m for all $m \in [m_0, +\infty]$,

$$\frac{d}{d\alpha}\mathsf{S}_m(\rho_\alpha)\bigg|_{\alpha=0} \leqslant \mathsf{S}_m(\nu) - \mathsf{S}_m(\rho),$$

Adding the three previous inequalities gives the result. $\qquad \square$

Next, we apply the previous proposition to characterize elements belonging to the subdifferential of E_m for $m_0 \leqslant m < +\infty$. Our proof generalizes Ambrosio, Gigli, and Savaré's characterization of the subdifferential of aggregation-diffusion energies to the case of nonconvex, singular interaction potentials satisfying hypotheses (LSC), (GF1)–(GF3) (cf. [2, Theorem 10.4.13]), and we defer its proof to appendix section 8.2.

Proposition 4.7. *Suppose K satisfies hypotheses (LSC), (GF1)–(GF3) and $m \in [m_0, +\infty)$. Then,*

$$(4.9) \qquad |\partial \mathsf{E}_m|(\rho) < +\infty \iff \begin{cases} \rho^m \in W^{1,1}(\mathbb{R}^d), \\ (\nabla K * \rho) + \frac{\nabla \rho^m}{\rho} \in \partial \mathsf{E}_m(\rho), \\ |\partial \mathsf{E}_m|(\rho) = \left\| (\nabla K * \rho) + \frac{\nabla \rho^m}{\rho} \right\|_{L^2(\rho)}. \end{cases}$$

In particular, for all $\rho \in D(|\partial \mathsf{E}_m|)$, $(\nabla K * \rho) + \frac{\nabla \rho^m}{\rho}$ is the unique element of the subdifferential of E_m at ρ with minimal $L^2(\rho)$-norm.

Remark 4.8 (division by ρ). For simplicity, we commit a small notational abuse in the above expression of the minimal element of the subdifferential: we divide by ρ, even though ρ may not be strictly positive. More precisely, let $w = (\nabla K * \rho) + \frac{\nabla \rho^m}{\rho}$ represent a function satisfying $(\nabla K * \rho)\rho + \nabla \rho^m = w\rho$ almost everywhere. This function is unique ρ-almost everywhere.

We now identify elements belonging to the subdifferential of E_∞, providing an upper bound on the metric slope of E_∞. (See Theorem 1.10 for the characterization of minimal elements of $\partial\mathsf{E}_\infty$ along gradient flows.)

Proposition 4.9. *If K satisfies hypotheses (LSC), (GF1)–(GF3) and $\rho \in D(\mathsf{E}_\infty)$,*

$$\nabla K * \rho \in \partial\mathsf{E}_\infty(\rho) \text{ and } \|\nabla K * \rho\|_{L^2(\rho)} \geqslant |\partial\mathsf{E}_\infty|(\rho).$$

Proof. It suffices to show $\nabla K * \rho \in \partial\mathsf{E}_\infty(\rho)$, as the second inequality then follows from Remark 2.1. By Proposition 4.6, it suffices to show for all $\nu \in D(\mathsf{E}_\infty)$,

(4.10) $\mathsf{E}_\infty(\nu) - \mathsf{E}_\infty(\rho)$

$$\geqslant \int \left\langle \nabla K * \rho, \mathbf{t}_\rho^\nu - \mathrm{Id} \right\rangle d\rho - f(\rho,\nu)\psi(d_W(\rho,\nu))d_W(\rho,\nu), \ \forall \nu \in D(\mathsf{E}_m),$$

By definition of E_∞ and the fact that $\nu, \rho \in D(\mathsf{E}_\infty)$, this is equivalent to

(4.11) $$\mathsf{K}(\nu) - \mathsf{K}(\rho) \geqslant \int \langle \nabla K * \rho, \mathbf{t}_\rho^\nu - \mathrm{Id}\rangle d\rho - f(\rho,\nu)\psi(d_W(\rho,\nu))d_W(\rho,\nu),$$

which is an immediate consequence of Proposition 4.5.   □

We apply the previous results to prove Theorem 1.7, which ensures that, for any initial data $\rho_0 \in D(\mathsf{E}_m)$, the gradient flow exists. It also provide sufficient conditions for the gradient flow to be unique.

Proof of Theorem 1.7. Existence follows from Proposition 4.2, [2, Corollary 11.1.8], and [2, Example 11.1.9]. Uniqueness follows from Proposition 4.5, the convexity of S_m for all $m \in (1, +\infty]$, and [42, Proposition 2.8, Theorem 3.12], with $\omega(s) = \sqrt{s\psi(s)}$. The correspondence between solutions of the aggregation-diffusion equation and gradient flows of E_m when $m < +\infty$ follows from the characterization of the minimal element of the subdifferential from Proposition 4.7 and [2, Corollary 11.1.8].   □

We conclude by proving that gradient flows of E_m satisfy an energy dissipation identity.

Proposition 4.10. *Suppose K satisfies hypotheses (LSC) and (GF1)–(GF3), $m \in [m_0, +\infty]$, and $\rho_m^{(0)} \in D(\mathsf{E}_m)$. Then if $\rho_m(t)$ is the gradient flow of E_m with initial data $\rho_m^{(0)}$, for all $T > 0$,*
  (i) $|\rho_m'|(t) = |\partial\mathsf{E}_m|(\rho_m(t)) < +\infty$ for a.e. $t \in [0,T]$;
  (ii) $\mathsf{E}_m(\rho_m(T)) + \int_0^T |\partial\mathsf{E}_m|^2(\rho_m(t))\,dt = \mathsf{E}_m(\rho_m^{(0)})$.

Proof. The result follow from the fact that the gradient flow is a curve of maximal slope for the strong upper gradient $|\partial\mathsf{E}_m|$; see Corollary 8.3, [2, Theorem 11.1.3], and [2, Remark 1.3.3]. □

5. Convergence of Minimizers

In this section, we prove our first main results: up to a sequence, minimizers of the energies E_m converge to minimizers of the energies E_∞ and these minimizers have uniformly bounded support. We begin by proving the Γ-convergence of E_m to E_∞.

Theorem 5.1 (Γ-convergence of E_m to E_∞)**.** *Suppose K satisfies (LSC). If $\rho_m \overset{*}{\rightharpoonup} \rho$, then*

$$\liminf_{m \to +\infty} \mathsf{E}_m(\rho_m) \geqslant \mathsf{E}_\infty(\rho).$$

Furthermore, for any $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, we have $\limsup_{m \to +\infty} \mathsf{E}_m(\rho) \leqslant \mathsf{E}_\infty(\rho)$.

Proof. Without loss of generality, we may assume that $\liminf_{n\to+\infty} \mathsf{E}_m(\rho_m) < +\infty$. Taking a subsequence ρ_m so that the lim inf is attained, we may also assume that $\sup_m \mathsf{E}_m(\rho_m) < +\infty$. Applying Proposition 4.1 and interpolation of $L^p(\mathbb{R}^d)$ norms for $1 < m_0 \leqslant m$, we obtain that $\sup_m \|\rho_m\|_{m_0} < +\infty$. Therefore, by Proposition 4.2, $\liminf_{m\to+\infty} \mathsf{K}(\rho_m) \geqslant \mathsf{K}(\rho)$. In particular, $\sup_m \mathsf{K}(\rho_m) > -\infty$.

It remains to show that $\liminf_{m\to+\infty} \mathsf{S}_m(\rho_m) \geqslant \mathsf{S}_\infty(\rho)$. Note that

$$\sup_m \mathsf{E}_m(\rho_m) < +\infty \quad \text{and} \quad \sup_m \mathsf{K}(\rho_m) > -\infty \implies \sup_m \mathsf{S}_m(\rho_m) < +\infty.$$

Therefore, there exists $C > 1$ so that $\|\rho_m\|_m \leqslant C^{1/m}(m-1)^{1/m} \leqslant C^{1/m}$ for $m \geqslant m_0$. By the interpolation of $L^p(\mathbb{R}^d)$ norms, for any $r \in [1, m]$,

$$\|\rho_m\|_r \leqslant \|\rho_m\|_m^{1-\theta} \leqslant C^{(1-\theta)/m} \leqslant C^{(1-\theta)/r} \leqslant C^{1/r},$$

where $0 \leqslant \theta \leqslant 1$ satisfies $1/r = \theta + (1-\theta)/m$. Since the $L^r(\mathbb{R}^d)$-norm is lower semicontinuous with respect to weak-* convergence, we obtain

$$\|\rho\|_r \leqslant \liminf_{m\to\infty} \|\rho_m\|_r \leqslant C^{1/r}.$$

Sending $r \to +\infty$ then yields

$$\|\rho\|_\infty \leqslant 1.$$

Therefore $\mathsf{S}_\infty(\rho) = 0$, and since S_m is positive, we have

$$\liminf_{m\to+\infty} \mathsf{S}_m(\rho_m) \geqslant \mathsf{S}_\infty(\rho).$$

We now turn to the lim sup inequality. Without loss of generality, we may assume $\mathsf{E}_\infty(\rho) < +\infty$, so $\|\rho\|_\infty \leqslant 1$. Applying Hölder's inequality gives $|\mathsf{S}_m(\rho)| \leqslant \frac{1}{m-1}\|\rho\|_\infty^{m-1} \to 0 = \mathsf{S}_\infty(\rho)$ as $m \to +\infty$, which gives the result. $\qquad\square$

We now prove that minimizers of E_m and E_∞ exist and are compactly supported.

Proposition 5.2 (existence of minimizers in $\mathcal{P}_2(\mathbb{R}^d)$)**.** *Suppose K satisfies (LSC) and (ATT). Then E_∞ and E_m admit compactly supported minimizers in $\mathcal{P}_2(\mathbb{R}^d)$ for all $m > \max\{m_0, 2\}$.*

Remark 5.3. Although it is possible to prove the existence of minimizers in the regime $1 < m < 2$, this requires additional assumptions in the hypothesis (ATT) (cf. hypothesis (K6) in [26]). Since we are interested in the large m regime we choose to prove the above theorem for $m > \max\{m_0, 2\}$.

Proof. If K satisfies (ATT)(i), then the existence of minimizers of E_∞ follows from the fact that the energy is decreasing under symmetric decreasing rearrangements of ρ (see e.g. [20, Proposition 3.1]). For E_m, existence of compactly supported minimizers is established in [26, Theorem 3.1 and Lemma 3.7].

If K satisfies (ATT)(ii), then the interaction potential K is strictly increasing in every coordinate outside of some fixed set, and the existence of a minimizer in $\mathcal{P}(\mathbb{R}^d)$ for the energy E_∞ over $\mathcal{P}(\mathbb{R}^d)$ follows simply by [20, Proposition 4.1]. The minimizer in $\mathcal{P}(\mathbb{R}^d)$ is in fact compactly supported by [20, Lemma 4.4], and therefore is in $\mathcal{P}_2(\mathbb{R}^d)$.

The existence of a minimizer in $\mathcal{P}(\mathbb{R}^d)$ for E_m also follows by using the growth of K given by (ATT)(ii), and by adapting the arguments in [65, Theorem 3.1]. This strong coercive behavior of K is sufficient to obtain the existence of a minimizer in $\mathcal{P}(\mathbb{R}^d)$ as the diffusion term is bounded from below (for $m > 1$); hence, can be controlled by the growth of K.

In order to conclude that the minimizer is indeed in $\mathcal{P}_2(\mathbb{R}^d)$ we need to show that it is compactly supported in $\mathbb{R}^d$. To this end, note that, if ρ minimizes E_m in $\mathcal{P}(\mathbb{R}^d)$ then a simple calculation shows that it satisfies the first-order variational inequality

$$(K * \rho)(x) \leqslant (K * \rho)(x) + \frac{m}{m-1}\rho^{m-1}(x) \leqslant \lambda$$

for some $\lambda \in \mathbb{R}$ and for all $x \in \operatorname{supp}\rho$. Note that

$$(K * \rho)(x) \geqslant \int_{|y| \leqslant R} K(x - y)\, d\rho(y) \geqslant C_R \inf\left\{K(z)\colon |z| > |x| - R\right\}$$

where $R > 0$ is chosen large enough so that $C_R := \rho(\{y\colon |y| < R\}) > 0$. Thus $\lim_{|x|\to\infty}(K * \rho)(x) \to \infty$; hence, $\left\{x \in \mathbb{R}^d\colon (K * \rho)(x) \leqslant \lambda\right\}$ is bounded, and so $\operatorname{supp}\rho$ is compact. $\qquad\square$

An important step in the proof of Theorem 1.3 is the compactness of a sequence of admissible measures whose E_m-energy is uniformly bounded. The main idea in proving such a compactness theorem is to utilize Lions' concentration compactness lemma [52] in order to show that any sequence of probability measures with uniformly bounded energy is tight up to translations. The proof of the following lemma follows by arguing as in [9] and [45].

Lemma 5.4 (Compactness in $\mathcal{P}(\mathbb{R}^d)$). *Suppose K satisfies (LSC) and (ATT). Let $\{\rho_m\}_{m>1} \subset \mathcal{P}(\mathbb{R}^d)$ be a sequence so that $\sup_m \mathsf{E}_m(\rho_m) < +\infty$. Then, up to translations, a subsequence of $\{\rho_m\}_{m>1}$ converges to a measure $\rho \in \mathcal{P}(\mathbb{R}^d)$ with respect to the weak-* topology.*

Now we turn to our convergence result.

Proof of Theorem 1.3. The convergence of minimizers is a classical consequence of the Γ-convergence result in Theorem 5.1, when the sequence of energies satisfies a sequential compactness property. Let $\{\rho_m\}_{m>1} \subset \mathcal{P}_2(\mathbb{R}^d)$ be a sequence of minimizers of E_m over $\mathcal{P}_2(\mathbb{R}^d)$. Then there exists $C > 0$ such that $\mathsf{E}_m(\rho_m) \leqslant C$ for $m > 1$ sufficiently large. Hence, by Lemma 5.4, there exists $\rho \in \mathcal{P}(\mathbb{R}^d)$ such that, up to a subsequence, $\rho_m \overset{*}{\rightharpoonup} \rho$ as $m \to +\infty$ in the weak-* topology of $\mathcal{P}(\mathbb{R}^d)$.

Now let $\nu \in P_2(\mathbb{R}^d)$ be arbitrary. Then

$$\mathsf{E}_\infty(\rho) \leqslant \liminf_{m\to+\infty} \mathsf{E}_m(\rho_m) \leqslant \liminf_{m\to+\infty} \mathsf{E}_m(\nu) = \mathsf{E}_\infty(\nu).$$

However, since minimizers of E_∞ are compactly supported we have that

$$\inf_{\mathcal{P}(\mathbb{R}^d)} \mathsf{E}_\infty = \inf_{\mathcal{P}_2(\mathbb{R}^d)} \mathsf{E}_\infty.$$

Therefore ρ minimizes E_∞ over $\mathcal{P}(\mathbb{R}^d)$, and since it is compactly supported, we have $\rho \in \mathcal{P}_2(\mathbb{R}^d)$. $\qquad\square$

Next we show that, for particular choices of K, a sequence of minimizers $\{\rho_m\}_{m>1}$ has compact support uniform in m. In order to establish this we adapt the arguments by Rein [63] to purely attractive interaction potentials, and follow the method by Frank and Lieb [46] to handle repulsive-attractive interactions. This allows us to conclude the sequence of minimizers converges in the 2-Wasserstein metric, despite fact that the compactness result, Lemma 5.4, holds only in $\mathcal{P}(\mathbb{R}^d)$. This gives Corollary 1.6.

Proof of Theorem 1.5. In order to prove this theorem for interaction potentials $K(x) = \frac{1}{p}|x|^p$, $-d < p < 0$, we proceed similarly to [63] where in the regime $d = 3$ and $p = -1$, Rein obtains a bound on the support of minimizers of E_m independent of the diffusion term. Let

$$I_M := \inf\left\{ \mathsf{E}_m(\rho) \colon \int \rho\, dx = M \right\}.$$

Using the scaling properties of the energy functional E_m under the transformations of the form $\rho(x) \mapsto l_1 \rho(l_2 x)$ it is easy to see (cf. [63, Lemma 3.5]) that $I_M < 0$ for all $M > 0$ by taking $l_1 = l_2^d$, and for $0 < \bar{M} \leqslant M$, we have

$$(5.1) \qquad\qquad I_{\bar{M}} \geqslant \left(\bar{M}/M \right)^{(2d+p)/d} I_M$$

by taking $l_1 = 1$ and $l_2 = (M/\bar{M})^{1/d}$.

Now, let $\rho \in L^1(\mathbb{R}^d) \cap L^m(\mathbb{R}^d)$ be any spherically symmetric, nonnegative function with $\|\rho\|_1 = 1$, and define $M_R := \int_{|x| \geqslant R} \rho\, dx$ for any $R > 0$. Then the splitting of the energy

$$\mathsf{E}_m(\rho) = \mathsf{E}_m(\rho \chi_{B_R}) + \mathsf{E}_m(\rho \chi_{B_R^c}) + \iint K(x-y)\chi_{B_R}(x)\chi_{B_R^c}(y)\, d\rho(x)d\rho(y),$$

combined with the estimate (which follows due to the spherical symmetry of ρ)

$$\left| \iint K(x-y)\chi_{B_R}(x)\chi_{B_R^c}(y)\, d\rho(x)d\rho(y) \right| \leqslant C(1-M_R)M_R R^p$$

implies that

$$\begin{aligned} \mathsf{E}_m(\rho) &\geqslant I_{M_R} + I_{1-M_R} - (1-M_R)M_R R^p \\ &\geqslant \left((1-M_R)^{(2d+p)/d} + M_R^{(2d+p)/d} \right) I_1 - (1-M_R)M_R R^p \\ &\geqslant \left(1 - \frac{2d+p}{d}(1-M_R)M_R \right) I_1 - (1-M_R)M_R R^p. \end{aligned}$$

where we have used (5.1) in the second line and Taylor's expansion in the third line. Defining

$$R_0 := -\frac{d}{(2d+p)I_1}$$

the above estimate becomes

$$(5.2) \qquad\qquad \mathsf{E}_m(\rho) \geqslant I_1 + \left(\frac{1}{R_0} - R^p \right)(1-M_R)M_R.$$

Take $R > R_0^{-1/p}$, and assume that for any spherically symmetric minimizing sequence $\{\rho_k\}_{k \in \mathbb{N}}$ for I_1 we have, up to a subsequence, that $\lim_{k \to \infty} \int_{|x| \geqslant R} \rho_k\, dx = M_R > 0$. Choose $R_k > R$ such that $M_k = \int_{|x| \geqslant R_k} \rho_k\, dx = 1/2 \int_{|x| \geqslant R} \rho_k\, dx$. Then, by (5.2),

$$\mathsf{E}_m(\rho_k) \geqslant I_1 + \left(\frac{1}{R_0} - R_k^p \right)(1-M_k)M_k \geqslant I_1 + \left(\frac{1}{R_0} - R^p \right)(1-M_k)M_k.$$

Sending $k \to \infty$, we get

$$I_1 \geqslant I_1 + \left(\frac{1}{R_0} - R^p \right)\left(1 - \frac{M_R}{2} \right)\frac{M_R}{2} > I_1;$$

a contradiction. Therefore, the minimizer of E_m, which is the weak-* limit of ρ_k, is supported in the ball of radius $R_0^{-1/p}$.

For interaction potentials in the power-law form, given by $K(x) = \frac{1}{q}|x|^q - \frac{1}{p}|x|^p$ with $-d < p < 0 < q$, we adapt the arguments by Frank and Lieb [46] to our case. Let $K^a(x) = \frac{1}{q}|x|^q$ and $K^r(x) = -\frac{1}{p}|x|^p$ denote the attractive and repulsive parts of the interaction potential, and K^a and K^r the corresponding interaction energies, respectively, so that

$$\mathsf{E}_m(\rho) = \mathsf{S}_m(\rho) + \mathsf{K}^a(\rho) + \mathsf{K}^r(\rho).$$

Let $\rho \in L^1(\mathbb{R}^d) \cap L^m(\mathbb{R}^d)$ be any nonnegative function with $\|\rho\|_1 = 1$. Then

$$(K * \rho)(x) \geqslant \int_{\mathbb{R}^d \backslash B_R(x)} K^a(x - y)\, d\rho(y) \geqslant \frac{R^q}{q}\left(1 - \int_{B_R(x)} \rho\, dy\right)$$

$$\geqslant \frac{R^q}{q}\left(1 - \sup_{a \in \mathbb{R}^d} \int_{B_R(a)} \rho\, dy\right).$$

Together with the positivity of the diffusion term, this implies that

$$\mathsf{E}_m(\rho) \geqslant \frac{1}{2}\int (K * \rho)\, d\rho \geqslant \frac{R^q}{2q}\left(1 - \sup_{a \in \mathbb{R}^d} \int_{B_R(a)} \rho\, dy\right).$$

Therefore we get

$$(5.3) \qquad \sup_{a \in \mathbb{R}^d} \int_{B_R(a)} \rho\, dy \geqslant 1 - \frac{2q\mathsf{E}_m(\rho)}{R^q}.$$

Now let ρ be a minimizer of the energy E_m, and let $x \in \mathbb{R}^d$ be a given Lebesgue point of ρ with $\rho(x) > 0$. Let $r > 0$ be arbitrary, and define

$$\tilde{\rho}(y) := \rho(y/l_r)\chi_{B_r^c}(y/l_r) \qquad \text{with} \qquad l_r := \left(\int_{\mathbb{R}^d} \chi_{B_r^c}\, d\rho\right)^{-1/d}$$

so that $\|\tilde{\rho}\|_1 = 1$. We now suppress the dependence on x and denote by B_r the ball of radius r centered at $x \in \mathbb{R}^d$. Since $\chi_{B_r} + \chi_{B_r^c} \equiv 1$, we have

$$\mathsf{E}_m(\tilde{\rho}) = l_r^{2d+q}\mathsf{K}^a(\rho\chi_{B_r^c}) + l_r^{2d+p}\mathsf{K}^r(\rho\chi_{B_r^c}) + \frac{l_r^d}{m-1}\int \rho^m\chi_{B_r^c}\, dy$$

$$\leqslant l_r^{2d+q}\left(\mathsf{K}^a(\rho) - \int_{B_r} K^a * \rho\, d\rho(y) + \mathsf{K}^a(\rho\chi_{B_r})\right)$$

$$+ l_r^{2d+p}\left(\mathsf{K}^r(\rho) - \int_{B_r} K^r * \rho\, d\rho(y) + \mathsf{K}^r(\rho\chi_{B_r})\right) + l_r^d\left(\mathsf{S}_m(\rho) - \mathsf{S}_m(\rho\chi_{B_r})\right).$$

Since $\mathsf{E}_m(\rho) \leqslant \mathsf{E}_m(\tilde{\rho})$ and $l_r \geqslant 1$, we have

$$(5.4) \qquad \begin{aligned}\int_{B_r} K * \rho\, d\rho(y) &\leqslant \int_{B_r} K * \rho\, d\rho(y) + \frac{1}{m-1}\int_{B_r} \rho^m\, dy\\ &\leqslant \left(l_r^{2d+q} - 1\right)\mathsf{K}^a(\rho) + \left(l_r^{2d+p} - 1\right)\mathsf{K}^r(\rho) + \left(l_r^d - 1\right)\mathsf{S}_m(\rho)\\ &\qquad + l_r^{2d+q}\,\mathsf{K}^a(\rho\chi_{B_r}) + l_r^{2d+p}\,\mathsf{K}^r(\rho\chi_{B_r}).\end{aligned}$$

First, we will estimate the last two terms and show that they are of order $o(r^d)$. Since $-d < p < 0$, we have, by the Hardy-Littlewood-Sobolev inequality,

$$(5.5) \qquad \mathsf{K}^r(\rho\chi_{B_r}) \leqslant C\|\rho\|_{L^1(B_r)}\|\rho\|_{L^{d/(d+p)}(B_r)}.$$

Using interpolation of $L^p(\mathbb{R}^d)$ norms, for $m \geqslant d/(d+p)$,

$$\|\rho\|_{L^{d/(d+p)}(B_r)} \leqslant \|\rho\|_{L^1(B_r)}^{\theta}\|\rho\|_{L^m(B_r)}^{1-\theta} \leqslant C\|\rho\|_{L^1(B_r)}^{\theta}$$

where

$$\theta = \frac{\frac{d+p}{d} - \frac{1}{m}}{1 - \frac{1}{m}}.$$

On the other hand, by Hölder's inequality, $\|\rho\|_{L^1(B_r)} \leqslant C\|\rho\|_{L^m(\mathbb{R}^d)}r^{d-d/m}$. Combining these, (5.5) implies

$$(5.6) \qquad \mathsf{K}^r(\rho\chi_{B_r}) \leqslant C\|\rho\|_{L^1(B_r)}^{\theta+1} \leqslant Cr^{2d(m-1)/m+p}.$$

Since the attractive part of the potential is strictly increasing a direct calculation shows that

$$(5.7) \qquad \mathsf{K}^a(\rho\chi_{B_r}) \leqslant \frac{1}{2q}\iint(K^a\chi_{B_{2r}(0)})(y-y')\,d\rho(y)d\rho(y') \leqslant Cr^{d+q},$$

where in the last step we use Young's inequality, and the embedding of $L^m(B_{2r}(0))$ into $L^2(B_{2r}(0))$ for $m \geqslant 2$.

Note that, since $-d < p$, we have $(d-p)/(2d) < 1$. Therefore, for m sufficiently large, $(m-1)/m > (d-p)/(2d)$, and $2d(m-1)/m + p - d > 0$. By the estimate (5.6),

$$\frac{1}{|B_r|}\mathsf{K}^r(\rho\chi_{B_r}) \leqslant Cr^{2d(m-1)/m+p-d} \xrightarrow{r\to 0} 0.$$

Moreover, the estimate (5.7) implies that

$$\frac{1}{|B_r|}\mathsf{K}^a(\rho\chi_{B_r}) \leqslant Cr^q \xrightarrow{r\to 0} 0.$$

Consequently, $\mathsf{K}^a(\rho\chi_{B_r}) + \mathsf{K}^r(\rho\chi_{B_r}) = o(r^d)$ as $r \to 0$; and since $l_r \to 1$ as $r \to 0$, the last two terms in (5.4) are of order $o(r^d)$, as well.

Since $x \in \mathbb{R}^d$ is a Lebesgue point of ρ, recalling the definition of l_r, we have

$$\frac{l_r^d - 1}{|B_r|} = \frac{1}{|B_r|(1 - \int_{B_r} \rho\,dy)}\int_{B_r}\rho\,dy \xrightarrow{r\to 0} \rho(x).$$

Also,

$$\frac{l_r^{2d+q} - 1}{l_r^d - 1} \to \frac{2d+q}{d} \quad\text{and}\quad \frac{l_r^d - 1}{l_r^d - 1} \to 1,$$

as $r \to 0$. Therefore, dividing both sides of (5.4) by $|B_r|$ and sending r to zero, we get

$$(5.8) \qquad K*\rho(x) \leqslant \frac{2d+q}{d}\mathsf{E}_m(\rho).$$

Now, let $R = (4q\mathsf{E}_m(\rho))^{1/q}$ so that $1/2 = 1 - 2qR^{-q}\mathsf{E}_m(\rho)$. By (5.3), there exists $a \in \mathbb{R}^d$ such that $\int_{B_R(a)} \rho \geqslant 1/2$. This implies that for every $y \in \mathbb{R}^d$ such that $|y-a| > (\sigma+1)R$ (with σ to be chosen shortly), we have

$$K*\rho(y) \geqslant \int_{B_R(a)} K^a(y-y')\,d\rho(y') \geqslant \frac{(|y-a|-R)^q}{q}\int_{B_R(a)}\rho\,dy' \geqslant \frac{(|y-a|-R)^q}{2q}$$
$$> \frac{\sigma^q R^q}{2q} = 2\sigma^q\mathsf{E}_m(\rho).$$

Let $\sigma := [(2d + q)/(2d)]^{1/q}$. Now, for $x \in \mathbb{R}^d$ is Lebesgue point of ρ such that $\rho(x) > 0$, combining (5.8) with the above estimate yields a contradiction if $|x - a| > (\sigma + 1)R$. Therefore $\rho(x) = 0$ for all $x \in \mathbb{R}^d$ with $|x - a| > (\sigma + 1)R$, and consequently,

$$\operatorname{diam} \operatorname{supp} \rho \leqslant 2(\sigma + 1)R = C(\mathsf{E}_m(\rho))^{1/q} \leqslant C'$$

for some constant $C' > 0$ independent of m when m is sufficiently large. $\qquad\square$

6. Convergence of Gradient Flows

In this section, we prove our main result on the convergence of gradient flows, Theorem 1.10. Throughout this section, we impose the following assumptions on our interaction potential K, diffusion exponent m, and the initial data of the gradient flows $\rho_m^{(0)}$.

Assumption 6.1 (Interaction potential, diffusion exponent, and initial data). *Suppose K satisfies hypotheses (LSC), (GF1)–(GF3), $m \geqslant m_0$, and $\sup_{m \geqslant m_0} \left(\mathsf{E}_m(\rho_m^{(0)}) + M_2(\rho_m^{(0)}) \right) < +\infty$.*

In the next proposition, we prove that several key quantities remain bounded along the gradient flow, uniformly in $m \geqslant m_0$. This plays a key role in our proof of Γ-convergence, since it provides weak compactness of the sequences $\nabla K * \rho_m$ and $\nabla \rho_m^m / \rho_m$ with respect to ρ_m, as well as weak compactness of ρ_m in arbitrarily large $L^p(\mathbb{R}^d)$ spaces.

Proposition 6.2 (Uniform bounds along gradient flow). *Fix $m < +\infty$. Suppose Assumption 6.1 holds and $\rho_m(t)$ is a gradient flow of E_m with initial data $\rho_m^{(0)}$. Then $\rho_m^m(t) \in W^{1,1}(\mathbb{R}^d)$ for a.e. $t > 0$, and for all $T > 0$,*

$$(6.1) \qquad \sup_{m \geqslant m_0} \int_0^T |\rho_m'|^2(t) + \|\nabla K * \rho_m\|_{L^2(\rho_m(t))}^2 + \|\nabla \rho_m^m(t)/\rho_m(t)\|_{L^2(\rho_m(t))}^2 \, dt < +\infty,$$

$$(6.2) \qquad \sup_{m \geqslant m_0, \, t \in [0,T]} \|\rho_m(t)\|_m < +\infty, \quad \text{and} \quad \sup_{m \geqslant \max\{m_0, d/2\}} \|\rho_m^m(t)\|_{L^2([0,T] \times \mathbb{R}^d)} < +\infty.$$

Proof. We begin by showing the first inequality in (6.2). By Propositions 4.1 and 4.10, there exists $C_r > 0$ so that

$$\|\rho_m(t)\|_m^{1 + \frac{1}{r}} \leqslant \mathsf{E}_m(\rho_m(t)) + C_r \leqslant \mathsf{E}_m(\rho_m^{(0)}) + C_r \text{ for all } t > 0.$$

By Assumption 6.1, $\sup_{m \geqslant m_0} \mathsf{E}_m(\rho_m^{(0)}) < +\infty$, which gives the result.

We now consider inequality (6.1). By Proposition 4.1, E_m is uniformly bounded below. By Assumption 6.1, $\mathsf{E}_m(\rho_m^{(0)})$ is uniformly bounded above. Combining this with Proposition 4.10,

$$(6.3) \qquad \sup_{m \geqslant m_0} \int_0^T |\rho_m'|^2(t) \, dt = \sup_{m \geqslant m_0} \int_0^T |\partial \mathsf{E}_m|^2(\rho_m(t)) \, dt < +\infty,$$

which is the first term in (6.1).

Next, we apply inequality (6.3) to obtain a uniform bound on the second moment of $\rho_m(t)$. By Hölder's inequality and the definition of the metric slope,

$$\sup_{m \geqslant m_0, \, t \in [0,T]} d_W(\rho_m^{(0)}, \rho_m(t)) \leqslant \sup_{m \geqslant m_0} \left(\int_0^T |\rho_m'|^2(t) \, dt \right)^{1/2} \sqrt{T} < +\infty.$$

Since Assumption 6.1 ensures $M_2(\rho_m^{(0)})$ is uniformly bounded, this implies

$$(6.4) \qquad \sup_{m \geqslant m_0, \, t \in [0,T]} M_2(\rho_m(t)) < +\infty.$$

We now turn to the second two terms in (6.1). By Proposition 4.7 and Assumption 6.1, for almost every $t > 0$, we have $\rho_m^m(t) \in W^{1,1}(\mathbb{R}^d)$ and

$$(6.5) \qquad |\partial \mathsf{E}_m|(\rho_m(t)) = \|\nabla K * \rho_m(t) + \nabla \rho_m^m(t)/\rho_m(t)\|_{L^2(\rho_m(t))}.$$

By hypothesis (GF1), the uniform bound on $\|\rho_m(t)\|_m$ from inequality (6.2), and the uniform bound on $M_2(\rho_m(t))$ from inequality (6.4), we have

$$(6.6) \qquad \sup_{m \geqslant m_0} \int_0^T \|\nabla K * \rho_m(t)\|_{L^2(\rho_m(t))}^2 \, dt < +\infty,$$

which is the second term in (6.1). Then combining equations (6.3), (6.5), and (6.6) with the triangle inequality gives

$$(6.7) \qquad \sup_{m \geqslant m_0} \int_0^T \|\nabla \rho_m^m(t)/\rho_m(t)\|_{L^2(\rho_m(t))}^2 \, dt < +\infty,$$

which is the third term in (6.1). This completes the proof of inequality (6.1).

We finally consider the second term in inequality (6.2). We proceed by using inequality (6.7) and our uniform bound on $\|\rho_m(t)\|_m$ to obtain improved estimates on ρ_m. Since $\rho_m^m(t) \in W^{1,1}(\mathbb{R}^d)$ for almost every $t > 0$, for any $\xi \in C_c^\infty([0,T] \times \mathbb{R}^d)$, there exists $C' > 0$ so that

$$
\begin{aligned}
\int_0^T \int \langle \nabla \rho_m^m(t), \xi(t) \rangle \, dx dt &= \int_0^T \int \langle \nabla \rho_m^m(t)/\rho_m(t), \xi(t) \rangle \, d\rho_m(t) dt \\
&\leqslant \int_0^T \|\nabla \rho_m^m(t)/\rho_m(t)\|_{L^2(\rho_m(t))} \|\xi(t)\|_{L^2(\rho_m(t))} \, dt \\
&\leqslant \int_0^T \|\nabla \rho_m^m(t)/\rho_m(t)\|_{L^2(\rho_m(t))} \|\rho_m(t)\|_m^{1/2} \|\xi(t)\|_{2m/(m-1)} \, dt \\
&\leqslant C' \left(\int_0^T \|\xi(t)\|_{2m/(m-1)}^2 \, dt \right)^{1/2}.
\end{aligned}
$$

Thus,

$$(6.8) \qquad \sup_{m \geqslant m_0} \int_0^T \|\nabla \rho_m^m(t)\|_{2m/(m+1)}^2 \, dt < +\infty.$$

Since $\rho_m^m(t) \in W^{1,1}(\mathbb{R}^d)$, $\rho_m^m(t)$ vanishes at $+\infty$. Thus, because we have $2m/(m+1) < 2$, we may apply the Sobolev embedding [50, Theorem 11.2], with

$$
q = \begin{cases} +\infty & \text{for } d = 1, \\ 2m^2 d/(m(d-2)+d) & \text{for } d \geqslant 2, \end{cases}
$$

which gives

$$(6.9) \qquad \|\rho_m(t)\|_q^m = \|\rho_m^m(t)\|_{q/m} \leqslant C'' \|\nabla \rho_m^m(t)\|_{2m/(m+1)}.$$

For $m \geqslant d/2$, we have $q \geqslant 2m$. Interpolating $L^p(\mathbb{R}^d)$-norms gives

$$
\|\rho_m^m(t)\|_2^2 = \|\rho_m(t)\|_{2m}^{2m} \leqslant \|\rho_m(t)\|_q^{q'(2m-1)}, \quad \text{for } q' = q/(q-1).
$$

Integrating in time and applying inequality (6.9),

$$(6.10) \qquad \int_0^T \|\rho_m^m(t)\|_2^2 \, dt \leqslant \int_0^T \|\rho_m(t)\|_q^{q'(2m-1)} \, dt \leqslant C'' \int_0^T \|\nabla \rho_m^m(t)\|_{2m/(m+1)}^{2q'/(2m)'} \, dt,$$

where $(2m)' := 2m/(2m-1)$. Since $q \geqslant 2m$, we have $q' \leqslant (2m)'$, and $s \mapsto s^{q'/(2m)'}$ is concave. Thus, applying Jensen's inequality to inequality (6.10) gives

$$\int_0^T \|\rho_m^m(t)\|_2^2 \, dt \leqslant C''T \left(\frac{1}{T} \int_0^T \|\nabla \rho_m^m(t)\|_{2m/(m+1)}^{2q'/(2m)'} \, dt \right)$$

$$\leqslant C''T \left(\frac{1}{T} \int_0^T \|\nabla \rho_m^m(t)\|_{2m/(m+1)}^2 \, dt \right)^{q'/(2m)'}.$$

By inequality (6.8), the right hand side is bounded uniformly in m, which gives the result. $\square$

We conclude with the proof of Theorem 1.10.

Proof of Theorem 1.10. First we show that there exists $\rho(t) \in \mathcal{P}_2(\mathbb{R}^d)$ so that, up to a subsequence, $\rho_m(t) \overset{*}{\rightharpoonup} \rho(t)$ for all $t \in [0, T]$. By definition of the metric derivative and Proposition 6.2, there exists $C' > 0$ so that for all $0 \leqslant s \leqslant t \leqslant T$ and $m \geqslant m_0$,

$$(6.11) \qquad d_W(\rho_m(s), \rho_m(t)) \leqslant \int_s^t |\rho_m'|(r) \, dr \leqslant \sqrt{t-s} \left(\int_s^t |\rho_m'|^2(t) \, dt \right)^{1/2} \leqslant C'\sqrt{t-s}.$$

In particular, taking $s = 0$ and recalling that $\sup_{m \geqslant m_0} M_2(\rho_m(0)) < +\infty$, we see that $\{\rho_m(t)\}_{m \geqslant m_0, t \in [0,T]}$ is uniformly bounded in $\mathcal{P}_2(\mathbb{R}^d)$, hence $\rho_m(t)$ is sequentially compact in the b-Wasserstein metric for all $1 \leqslant b < 2$ and $t \in [0, T]$ [2, Proposition 7.1.5]. Take $b = 2 - \epsilon$ for $\epsilon \in (0, 1)$ as in hypothesis (GF3). Then using the equicontinuity from inequality (6.11), the generalized Arzela-Ascoli/Aubin-Lions theorem [2, Proposition 3.3.1] implies that there exists $\rho(t) \in \mathcal{P}_2(\mathbb{R}^d)$ so that, up to a subsequence,

$$(6.12) \qquad \lim_{m \to +\infty} d_{W_{2-\epsilon}}(\rho_m(t), \rho(t)) = 0 \text{ for all } t \in [0, T].$$

In particular, up to a subsequence, we have $\rho_m(t) \overset{*}{\rightharpoonup} \rho(t)$ for all $t \in [0, T]$.

It remains to verify criteria (i)–(iii) of Theorem 2.2 to conclude that $\rho(t)$ is a gradient flow of E_∞ and that the corresponding energies, local slopes, and metric derivatives converge as $m \to +\infty$.

Criterion (i) follows immediately from Theorem 5.1, and as a consequence, we conclude that $\|\rho(t)\|_\infty \leqslant 1$ for all $t > 0$. Criterion (ii) is proved in Lemma 2.3. Thus, it remains to show criterion (iii). By Fatou's lemma, it suffices to show that

$$(6.13) \qquad \liminf_{m \to +\infty} |\partial \mathsf{E}_m|(\rho_m(s)) \geqslant |\partial \mathsf{E}_\infty|(\rho(s)), \quad \text{ for a.e. } s \in [0, t].$$

By Proposition 6.2 and Fatou's lemma,

$$(6.14) \qquad \int_0^t \liminf_{m \to +\infty} I_m(s) \, ds \leqslant \liminf_{m \to +\infty} \int_0^t I_m(s) \, ds < +\infty$$

$$\text{for } I_m(s) := |\partial \mathsf{E}_m|^2(\rho_m(s)) + \|v_m^1(s)\|_{L^2(\rho_m(s))}^2 + \|v_m^2(s)\|_{L^2(\rho_m(s))}^2 + \|\rho_m^m(s)\|_2^2$$

$$v_m^1(s) := (\nabla K * \rho_m)(s) \text{ and } v_m^2(s) := \nabla \rho_m^m(s)/\rho_m(s).$$

In particular, $\liminf_{m \to +\infty} I_m(s) < +\infty$ for almost every $s \in [0, T]$. Fix such an $s \in [0, T]$. We will now show that (6.13) holds, which completes the proof. As $s \in [0, T]$ is fixed, in what follows, we will suppress the dependence on time to ease notation.

Up to a subsequence, we have

$$\liminf_{m \to +\infty} I_m = \lim_{m \to +\infty} I_m < +\infty,$$

so that I_m is bounded uniformly in m. By Proposition 4.7, this ensures $\rho_m \in W^{1,1}(\mathbb{R}^d)$ and

$$(6.15) \qquad |\partial \mathsf{E}_m|(\rho_m) = \|v_m^1 + v_m^2\|_{L^2(\rho_m)}.$$

Since $\|v_m^1\|_{L^2(\rho_m)}$ and $\|v_m^2\|_{L^2(\rho_m)}$ are bounded uniformly in m, up to another subsequence, there exist $v^1, v^2 \in L^2(\rho)$ so that for all $f \in C_c^\infty(\mathbb{R}^d)$ [2, Theorem 5.4.4],

$$(6.16) \qquad \lim_{m \to +\infty} \int f v_m^1 \, d\rho_m = \int f v^1 \, d\rho, \quad \lim_{m \to +\infty} \int f v_m^2 \, d\rho_m = \int f v^2 \, d\rho,$$

and

$$(6.17) \qquad \liminf_{m \to +\infty} |\partial \mathsf{E}_m|(\rho_m) = \lim_{m \to +\infty} \|v_m^1 + v_m^2\|_{L^2(\rho_m)} \geqslant \|v^1 + v^2\|_{L^2(\rho)}.$$

By Remark 2.1, to complete the proof it suffices to show that $v^1 + v^2 \in \partial \mathsf{E}_\infty(\rho)$.

First, we will show that $v^1 = \nabla K * \rho$ ρ-almost everywhere. For all $f \in C_c^\infty(\mathbb{R}^d)$,

$$\left| \int f \left[v^1 - (\nabla K * \rho)\right] d\rho \right| = \lim_{m \to +\infty} \left| \int f \left[(\nabla K * \rho_m) \, d\rho_m - (\nabla K * \rho) \, d\rho\right] \right|$$

$$\leqslant \lim_{m \to +\infty} \left| \int f \left[(\nabla K * \rho) \, d\rho_m - (\nabla K * \rho) d\rho\right] \right| + \left| \int f \left[(\nabla K * \rho_m) - (\nabla K * \rho)\right] d\rho_m \right|$$

$$= \lim_{m \to +\infty} A_m + B_m.$$

By hypothesis (GF2), $(\nabla K * \rho)(x)$ is continuous in x. Thus $\rho_m \overset{*}{\rightharpoonup} \rho$ implies $\lim_{m \to +\infty} A_m = 0$. We now consider B_m. By Hölder's inequality and hypothesis (GF3), there exists $\epsilon > 0$ so

$$B_m \leqslant \|f\|_\infty \|\nabla K * (\rho_m - \rho)\|_{L^2(\rho_m)} \leqslant C\|f\|_\infty \psi(d_{W_{2-\epsilon}}(\rho_m, \rho)).$$

Since $\lim_{m \to +\infty} d_{W_{2-\epsilon}}(\rho_m, \rho) = 0$, $\lim_{m \to +\infty} B_m = 0$ and $v^1 = \nabla K * \rho$, ρ-almost everywhere.

By Proposition 4.9, $v^1 = \nabla K * \rho \in \partial \mathsf{E}_\infty(\rho)$. Thus, by the definition of the subdifferential (2.7), to complete our proof that $v^1 + v^2 \in \partial \mathsf{E}_\infty$, it suffices to show that

$$(6.18) \qquad \int \langle v^2, \mathbf{t}_\rho^\nu - \mathrm{Id} \rangle \, d\rho \leqslant 0, \quad \forall \nu \in D(\mathsf{E}_\infty).$$

We claim that

(6.19) $v^2 \rho = \nabla \sigma$ for $\sigma \in H^1(\mathbb{R}^d)$ satisfying $\sigma \geqslant 0$ and $\sigma = 0$ almost everywhere on $\{\rho < 1\}$.

We now show that this implies (6.18). Fix $\nu \in D(\mathsf{E}_\infty)$ and let $\mathbf{t}_\rho^\nu$ be the 2-Wasserstein optimal transport map from ρ to ν. Define $\mathbf{t}_\alpha = (1 - \alpha)\,\mathrm{Id} + \alpha \mathbf{t}_\rho^\nu$, $\alpha \in [0,1]$, so $\rho_\alpha = \mathbf{t}_{\alpha \#}\rho$ is the geodesic from ρ to ν. Since $\rho, \nu \in D(\mathsf{E}_\infty)$, $\|\rho_\alpha\|_\infty \leqslant \max\{\|\rho\|_\infty, \|\nu\|_\infty\} \leqslant 1$ for all $\alpha \in [0,1]$. The geodesic ρ_α satisfies the following weak form of the continuity equation for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ (cf. [2, equation (8.1.4)], [36])

$$(6.20) \qquad \int \varphi \, d\rho_\alpha - \int \varphi \, d\rho - \int_0^\alpha \int \langle \nabla \varphi, \mathbf{t}_\rho^\nu \circ \mathbf{t}_\beta^{-1} - \mathbf{t}_\beta^{-1} \rangle \, d\rho_\beta d\beta = 0.$$

Note that, for all $\beta \in [0,1]$, we have $\|\rho_\beta\|_2 \leqslant \|\rho_\beta\|_1^{1/2}\|\rho_\beta\|_\infty^{1/2} \leqslant 1$ and

$$\|(\mathbf{t}_\rho^\nu \circ \mathbf{t}_\beta^{-1} - \mathbf{t}_\beta^{-1})\rho_\beta\|_2 \leqslant \|\mathbf{t}_\rho^\nu \circ \mathbf{t}_\beta^{-1} - \mathbf{t}_\beta^{-1}\|_{L^2(\rho_\beta)}\|\rho_\beta\|_\infty^{1/2} \leqslant W_2(\rho, \nu).$$

Furthermore, we also have $\lim_{\beta \to 0}(\mathbf{t}_\rho^\nu \circ \mathbf{t}_\beta^{-1} - \mathbf{t}_\beta^{-1})\rho_\beta = (\mathbf{t}_\rho^\nu - \mathrm{Id})\rho$ in distribution. Finally, since $(\mathbf{t}_\rho^\nu \circ \mathbf{t}_\beta^{-1} - \mathbf{t}_\beta^{-1})\rho_\beta$ is uniformly bounded in $L^2(\mathbb{R}^d)$ for all $\beta \in [0,1]$, compactness with respect to the weak $L^2(\mathbb{R}^d)$ topology and uniqueness of limits implies

$$(6.21) \qquad \text{as } \beta \to 0, \quad (\mathbf{t}_\rho^\nu \circ \mathbf{t}_\beta^{-1} - \mathbf{t}_\beta^{-1})\rho_\beta \rightharpoonup (\mathbf{t}_\rho^\nu - \mathrm{Id})\rho \quad \text{in } L^2(\mathbb{R}^d).$$

Since $\sigma \in H^1(\mathbb{R}^d)$, approximating it by a sequence in $C_c^\infty(\mathbb{R}^d)$ and applying equation (6.20),

$$\int \sigma \, d\rho_\alpha - \int \sigma \, d\rho - \int_0^\alpha \int \langle \nabla\sigma, \mathbf{t}_\rho^\nu \circ \mathbf{t}_\beta^{-1} - \mathbf{t}_\beta^{-1} \rangle \, d\rho_\beta d\beta = 0.$$

As $\rho = 1$ wherever $\sigma \neq 0$, this is equivalent to

$$\frac{1}{\alpha} \int \sigma(\rho_\alpha - 1) \, dx = \frac{1}{\alpha} \int_0^\alpha \int \langle \nabla\sigma, \mathbf{t}_\rho^\nu \circ \mathbf{t}_\beta^{-1} - \mathbf{t}_\beta^{-1} \rangle \, d\rho_\beta d\beta.$$

Since $\sigma \geqslant 0$ and $\rho_\alpha \leqslant 1$, this implies that the left hand side is nonpositive for all $\alpha \in (0,1)$. Thus, sending $\alpha \to 0$ and using (6.21) gives

$$0 \geqslant \limsup_{\alpha \to 0} \frac{1}{\alpha} \int_0^\alpha \int \langle \nabla\sigma, \mathbf{t}_{\rho_\beta}^\nu - \mathbf{t}_{\rho_\beta}^\rho \rangle \, d\rho_\beta d\beta = \int \langle \nabla\sigma, \mathbf{t}_\rho^\nu - \mathrm{Id} \rangle \, d\rho.$$

Since $\rho \geqslant 0$, the integrand is nonpositive ρ-almost everywhere. Since $\rho \leqslant 1$, we obtain

$$\int_{\mathbb{R}^d} \langle \nabla\sigma, \mathbf{t}_\rho^\nu - \mathrm{Id} \rangle \, dx \leqslant \int \langle \nabla\sigma, \mathbf{t}_\rho^\nu - \mathrm{Id} \rangle \, d\rho \leqslant 0,$$

which shows (6.18).

It remains to show that the claim in equation (6.19). Since I_m is bounded uniformly in m, so is $\|\rho_m^m\|_2$, and there exists $\sigma \in L^2(\mathbb{R}^d)$ with $\sigma \geqslant 0$ so that, up to a subsequence, $\rho_m^m \rightharpoonup \sigma$ in $L^2(\mathbb{R}^d)$. Combining this fact with (6.16), we have that for all $f \in C_c^\infty(\mathbb{R}^d)$,

$$-\int \nabla f \sigma \, dx = -\lim_{m \to +\infty} \int \nabla f \rho_m^m \, dx = \lim_{m \to +\infty} \int f \nabla\rho_m^m \, dx = \lim_{m \to +\infty} \int f v_m^2 \, d\rho_m = \int f v^2 \, d\rho.$$

Since $\|v^2\rho\|_2 \leqslant \|v^2\|_{L^2(\rho)} \|\rho\|_\infty^{1/2} < +\infty$, we have that $v^2\rho = \sigma \in H^1(\mathbb{R}^d)$.

We conclude by showing that $\sigma = 0$ almost everywhere on $\{\rho < 1\}$, which is equivalent to

$$\text{(6.22)} \qquad\qquad\qquad\qquad \int \sigma(\rho - 1) \, dx = 0.$$

By Hölder's inequality and the uniform bound on I_m,

$$\sup_m \|\nabla\rho_m^m\|_1 \leqslant \sup_m \|\nabla\rho_m^m(t)/\rho_m\|_{L^2(\rho_m)} \|\rho_m\|_1^{1/2} = \sup_m \|v_m^2\|_{L^2(\rho_m)} < +\infty.$$

Fix $R > 0$ and let η_R be a smooth, radially decreasing cutoff function,

$$\eta_R(x) = \eta(x/R) \text{ for } \eta \in C^\infty(\mathbb{R}^d) \text{ such that } \eta(x) \equiv 1 \text{ for } |x| \leqslant \frac{1}{2} \text{ and } \eta(x) \equiv 0 \text{ for } |x| > 1.$$

Then we have

$$\|\nabla(\rho_m^m \eta_R)\|_1 \leqslant \|\nabla\rho_m^m\|_1 + \frac{1}{R}\|\nabla\eta\|_2 \|\rho_m^m\|_2, \quad \|\rho_m^m \eta_R\|_1 \leqslant \|\rho_m^m\|_2 \|\eta_R\|_2, \quad \text{and} \quad \|\rho_m^m \eta_R\|_2 \leqslant \|\rho_m^m\|_2$$

each of which is bounded uniformly in m. By Rellich-Kondrachov (cf. [50, Theorem 13.32]), there exists a subsequence so that $\rho_m^m \eta_R \to \sigma\eta_R$ strongly in $L^1(\mathbb{R}^d)$. Since $\rho_m^m \eta_R$ is uniformly bounded in $L^2(\mathbb{R}^d)$, we also obtain $\rho_m^m \eta_R \to \sigma\eta_R$ strongly in $L^p(\mathbb{R}^d)$ for all $1 \leqslant p < 2$.

Similarly, by interpolating $L^p(\mathbb{R}^d)$ norms, for any $p' > 2$ and $m \geqslant p'/2$,

$$\|\rho_m\|_{p'} \leqslant 1 + \|\rho_m\|_{2m},$$

which is bounded uniformly in m. Thus, up to a subsequence, $\rho_m \rightharpoonup \rho$ weakly in $L^{p'}(\mathbb{R}^d)$. Combining these two facts, we obtain,

$$\text{(6.23)} \qquad\qquad\qquad \lim_{m \to +\infty} \int \eta_R \rho_m^m(\rho_m - 1) \, dx \to \int \eta_R \sigma(\rho - 1) \, dx$$

Rewriting the left hand side of (6.23),

$$\int \eta_R \rho_m^m (\rho_m - 1)\, dx = \int \eta_R (\rho_m^m - \rho_m^{m-1})\, d\rho_m = \int \eta_R (\rho_m^m - \rho_m^{m(1-1/m)})\, d\rho_m$$

$$= \int \eta_R (\psi(\rho_m^m, 0) - \psi(\rho_m^m, b))\, d\rho_m,$$

where $\psi(s, a) := s^{(1-a)}$ and $b = 1/m$. By Lemma 8.5, we may control the right hand side by

$$\frac{1}{m} \int \eta_R \rho_m \left| 1 + \rho_m^{2m-1} \right|\, dx \leqslant \frac{1}{m} \int \eta_R\, d\rho_m + \frac{1}{m} \int \eta_R\, \rho_m^{2m}\, dx$$

which goes to zero uniformly as $m \to \infty$. Thus, for all $R > 0$,

$$\int \eta_R \sigma(\rho - 1)\, dx = 0.$$

Sending $R \to +\infty$ via the monotone convergence theorem, we obtain (6.22).

This concludes the proof of the criteria from Theorem 2.2. In particular, we have

$$|\partial \mathsf{E}_m|(\rho_m) \to |\partial \mathsf{E}_\infty|(\rho) \text{ in } L^2([0, T]).$$

Integrating inequality (6.17), we obtain that for all $f \in C^\infty([0, T])$ with $f \geqslant 0$,

$$\int_0^T |\partial \mathsf{E}_\infty(\rho(t))| f(t)\, dt = \liminf_{m \to +\infty} \int_0^T |\partial \mathsf{E}_m|(\rho_m(t)) f(t)\, dt$$

$$\geqslant \int_0^T \lim_{m \to +\infty} \|v_m^1(t) + v_m^2(t)\|_{L^2(\rho_m(t))} f(t)\, dt$$

$$\geqslant \int_0^T \|v^1(t) + v^2(t)\|_{L^2(\rho(t))} f(t)\, dt.$$

This gives $|\partial \mathsf{E}_\infty(\rho(t))| \geqslant \|v^1(t) + v^2(t)\|_{L^2(\rho(t))}$ for almost every $t \in [0, T]$. On the other hand, by Remark 2.1 and the fact that $v^1(t) + v^2(t) \in \partial \mathsf{E}_\infty(\rho(t))$, we have the opposite inequality. Therefore, equality holds and

$$v^1(t) + v^2(t) = \nabla K * \rho(t) + \nabla \sigma(t)/\rho(t)$$

is the element of $\partial E_\infty(\rho(t))$ with minimal $L^2(\rho(t))$ norm. $\qquad\qquad\qquad\square$

7. Numerical Results

In this section, we apply our theoretical results on the slow diffusion limit to develop a numerical method to simulate gradient flows and minimizers of the constrained interaction energy E_∞. For $m \geqslant m_0$ and $\rho_m^{(0)} \in D(\mathsf{E}_m)$, Theorem 1.7 ensures that $\rho_m(t)$ solves an aggregation-diffusion equation

$$(7.1) \qquad\qquad \partial_t \rho_m - \nabla \cdot ((\nabla K * \rho_m)\rho_m) = \Delta \rho_m^m, \quad \rho_m(0) = \rho_m^{(0)}$$

if and only if it is a gradient flow of E_m. By Theorem 1.10, gradient flows of E_m converge, up to a subsequence, to a gradient flow of E_∞. Thus, one may approximate the dynamics of gradient flows of E_∞ by numerically simulating solutions of equation (7.1) for m large. By Theorem 1.3, minimizers of E_m converge, up to a subsequence and translation, to minimizers of E_∞. Thus, to numerically approximate minimizers of E_∞, one may simulate gradient flows of E_m for m large, in the long time limit.

We now give several examples illustrating this approach, computed using Carrillo, Patacchini, and the first author's blob method for diffusion [35]. For present purposes, we merely consider simulations in one dimension, though our method extends naturally to all dimensions

$d \geqslant 1$. As the primary goal of present work is the rigorous analysis of the slow diffusion limit, we defer a more comprehensive numerical study to future work. Throughout, we take the regularization parameter ϵ in the blob method for diffusion to depend on the spatial grid spacing h according to $\epsilon = h^{.999}$. The initial data for our simulations is either patch initial data,

$$(7.2) \qquad\qquad\qquad \rho_m^{(0)} = \chi_\Omega \text{ for some } \Omega \subseteq \mathbb{R}^d,$$

or Barenblatt profiles, for $m_* > 1$ and $\tau > 0$

$$(7.3) \quad \rho_m^{(0)}(x) = \tau^{-d\beta}(K - \kappa\tau^{-2\beta}|x|^2)_+^{1/(m_*-1)}, \quad \beta = \frac{1}{2 + d(m_* - 1)}, \quad \kappa = \frac{\beta}{2}\left(\frac{m_* - 1}{m_*}\right),$$

with $K = K(m_*, d) > 0$ chosen so that $\int \rho_m^{(0)}\, dx = 1$.

In Figure 1, we simulate minimizers of E_m to study how the support of minimizers depends on the diffusion exponent m and the mass of the initial data $\int \rho_m^{(0)}\, dx$. This complements our theoretical result from Theorem 1.5 that the support of minimizers of E_m is uniformly bounded for m sufficiently large. For the purely attractive quadratic interaction potential $K(x) = 2|x|^2$, the size of the support of minimizers is decreasing in m for $\int \rho_m^{(0)}\, dx = 1$, constant for $\int \rho_m^{(0)}\, dx = 2$, and increasing in m for $\int \rho_m^{(0)}\, dx = 3$. This simulation demonstrates that monotonicity properties of the size of the support of minimizers strongly depend on the choice of interaction potential K and the mass of the initial data.

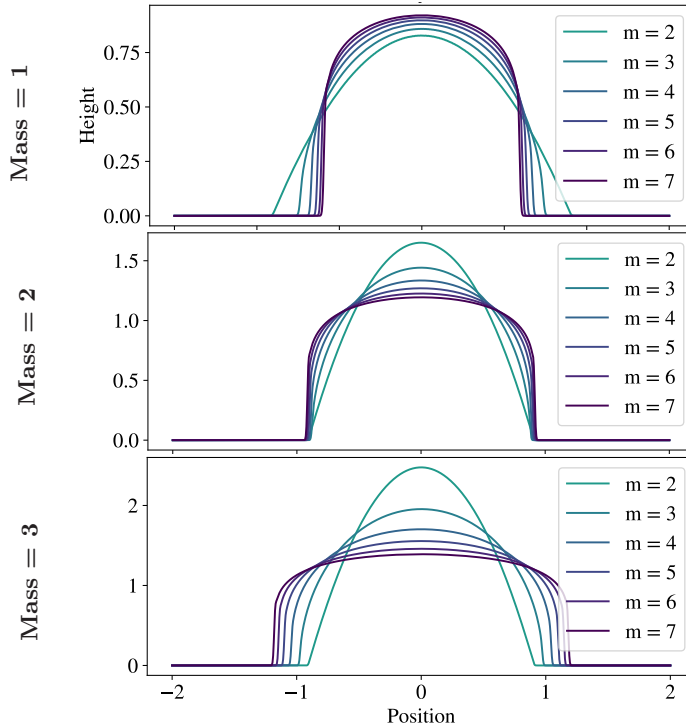**Equilibria of Aggregation-Diffusion Equation for Varying $m > 1$ and mass**

Figure 1. Minimizers of E_m, for $K(x) = 2|x|^2$ and varying m and $\int \rho_m\, dx$. Solutions of aggregation-diffusion equations are simulated at time $T = 10$ with spatial and temporal discretizations $h = 0.007$, $k = 10^{-4}$. The initial data is a constant multiple of a Barenblatt profile (equation (7.3), $m_* = 2$, $\tau = 0.15$).
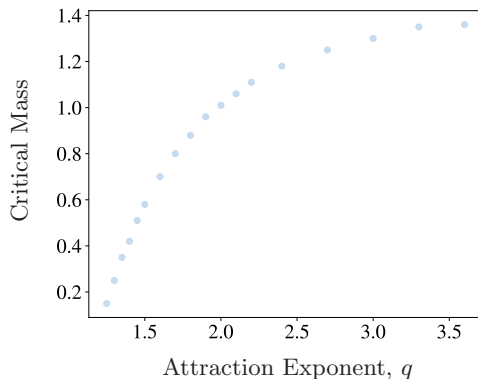
Critical Mass of Set-Valued Minimizers of E_∞ for Varying q

Figure 2. We approximate the critical mass that determines existence of set-valued minimizers of E_∞ by simulating solutions of aggregation-diffusion equations for $m = 800$, $K(x) = |x|^q/q - |x|$, $h = 0.004$, $k = 10^{-4}$, and various choices of initial data. For both initial data given by constant multiples of Barenblatt profiles (equation (7.3), $m_* = 2$, $\tau = 0.1$) and patch functions (equation (7.2), $\Omega = [-1, 1]$), we observe the existence of set valued minimizers occurs for the same value of mass. Above, we plot how this critical mass value depends on q.

Next, we consider gradient flows and minimizers of the constrained interaction energy E_∞ with repulsive-attractive power-law interaction potentials of the form

$$K(x) = |x|^q/q - |x|^p/p, \quad d - 2 \leqslant p \leqslant q.$$

We take the repulsion exponent $p = 1$ (the Newtonian singularity in one dimension) and allow the attraction exponent q to vary. We apply our numerical method for constrained interaction energies to explore open questions related to minimizers of E_∞, as described in the introduction.

In Figure 2, we investigate the value of the critical mass that determines existence versus nonexistence of set valued minimizers of E_∞. In particular, for initial data that is either a constant multiple of a Barenblatt profile or patch function, we observe that there exists a single value of the critical mass that separates existence and non-existence of set valued minimizers: in the notation from the introduction, we find $M_1 = M_2$ for all $q \geqslant 1$. We plot how the value of the critical mass depends on the attraction exponent $q \geqslant 1$.

In Figure 3, we approximate gradient flows of the constrained interaction energy E_∞ for various choices of attraction parameter $q \geqslant 1$, with initial data at the critical mass from Figure 2, i.e., the smallest value of mass for which solutions approach a set valued equilibrium. We contrast the behavior of initial data that is a constant multiple of a Barenblatt profile (equation (7.3), $m_* = 2$, $\tau = 1$)) with initial data that is a constant multiple of a patch function (equation (7.2), $\Omega = [-1, 1]$). In both cases, gradient flows converge to a characteristic function of height one on an interval centered at the origin. When $q < 2$, solutions initially reach height one at the center of mass of the density and then spread to become a characteristic function. When $q > 2$, solutions initially reach height one at the boundary of the support of the density and then "fill in" the interior to become a characteristic function.

Finally, in Figure 4, we compute minimizers of the interaction energy E_∞ for varying choices of attraction exponent $q \geqslant 1$ and masses up to and including the critical mass from Figure 2. These simulations appear to confirm the existence of an *intermediate phase* between the liquid and solid phase as the generic behavior for $q \neq 2$. In the notation of the introduction, these simulations suggest that $M_1^* < M_2^*$ for $q \neq 2$. In particular, we observe that minimizers of mass
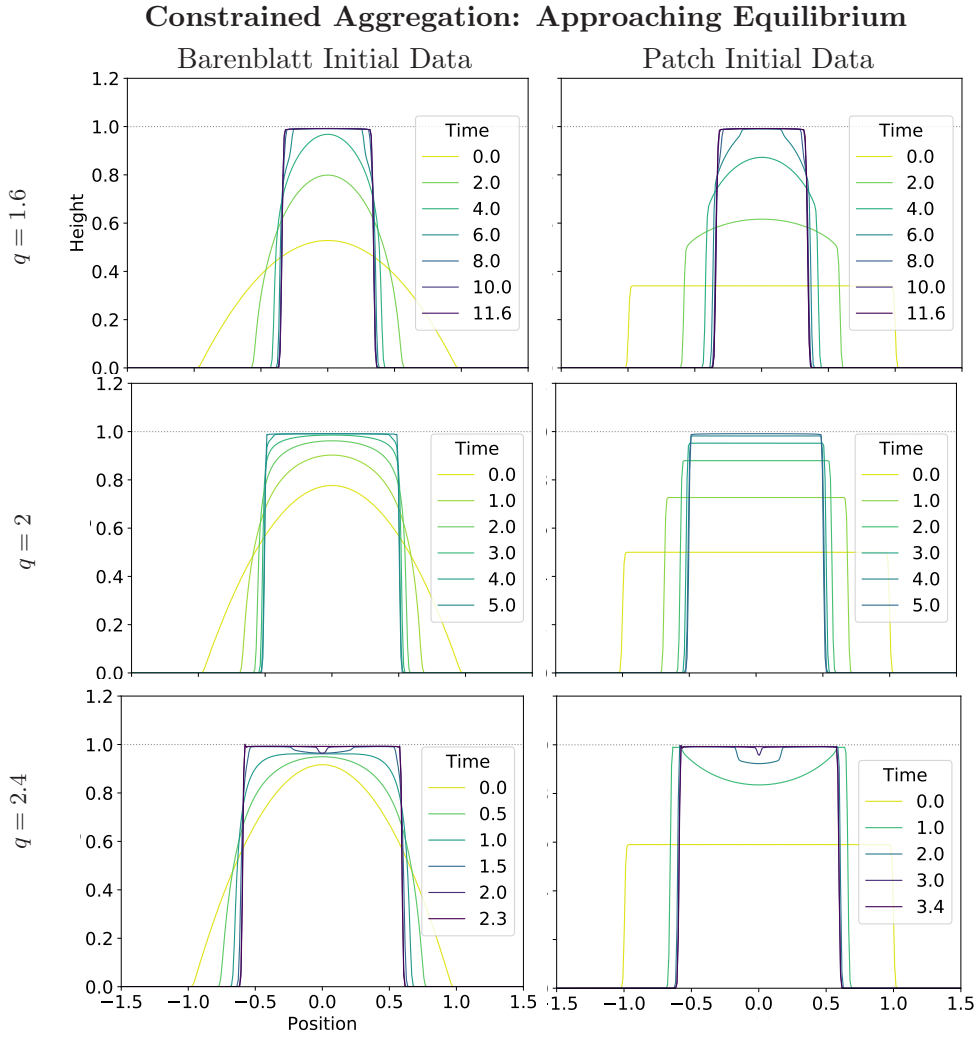
Figure 3. We approximate gradient flows of E_∞ by simulating solutions of aggregation-diffusion equations for $m = 800$, $K(x) = |x|^q/q - |x|$, $h = 0.004$, $k = 10^{-4}$, and initial data given by constant multiples of Barenblatt profiles (equation (7.3), $m_* = 2$, $\tau = 0.1$) and patch functions (equation (7.2), $\Omega = [-1, 1]$). We multiply the initial data by the desired mass for each q, i.e. mass 0.68, 1.00, and 1.18, for $q = 1.6$, 2.0, and 2.4.

M satisfy $|\{\rho = 1\}| \in (0, M)$ for all $M \in [0.36, 0.42]$ when $q = 1.4$ and for all $M \in [0.99, 1.19]$ for $q = 2.6$.

8. APPENDIX

8.1. **Regular Functionals.** Consider a functional $\mathsf{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ that is proper, lower semicontinuous, and satisfies $D(|\partial \mathsf{F}|) \subseteq \mathcal{P}_{2,ac}(\mathbb{R}^d)$. Then $\xi \in \partial \mathsf{F}(\mu)$ is a *strong subdifferential* if for all Borel measurable functions $\mathbf{t} : \mathbb{R}^d \to \mathbb{R}^d$ we have

$$\mathsf{F}(\mathbf{t}\#\mu) - \mathsf{F}(\mu) \geqslant \int \langle \xi, \mathbf{t} - \mathrm{Id} \rangle \, d\mu + o(\|\mathbf{t} - \mathrm{Id}\|_{L^2(\mu)}).$$

Following Ambrosio, Gigli, and Savaré [2, Definition 10.1.4], we define the notion of *regular functional* as follows.

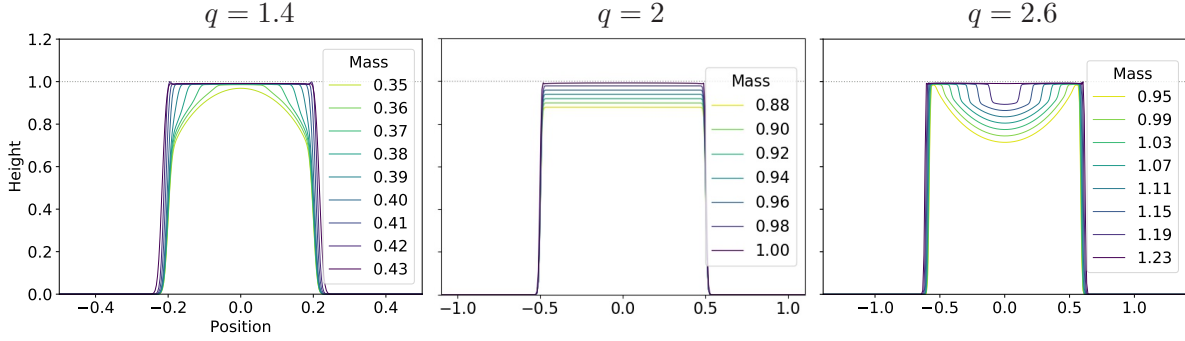## Constrained Aggregation: Equilibria for Varying Masses

Figure 4. We approximate minimizers of E_∞ by simulating solutions of aggregation-diffusion equations for $m = 800$, $K(x) = |x|^q/q - |x|$, $h = 0.004$, $k = 10^{-4}$, with initial data given by constant multiples of Barenblatt profiles (equation (7.3), $m_* = 2$, $\tau = 0.1$).

Definition 8.1. *Given* $\mathsf{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ *proper, lower semicontinuous, and satisfying* $D(|\partial \mathsf{F}|) \subseteq \mathcal{P}_{2,ac}(\mathbb{R}^d)$, F *is a* regular *functional if for any 2-Wasserstein convergent sequence* $\mu_n \to \mu$ *with strong subdifferentials* $\xi_n \in \partial \mathsf{F}(\mu_n)$ *satisfying*

(i) $\sup_n |\mathsf{F}(\mu_n)| < +\infty$
(ii) $\sup_n \|\xi_n\|_{L^2(\mu_n)} < +\infty$
(iii) *there exists* $\xi \in L^2(\mu)$ *so that* $\lim_{n \to +\infty} \int f \xi_n \, d\mu_n = \int f \xi \, d\mu$ *for all* $f \in C_c^\infty(\mathbb{R}^d)$,

we have $\lim_{n \to +\infty} \mathsf{F}(\mu_n) = \mathsf{F}(\mu)$ *and* $\xi \in \partial \mathsf{F}(\mu)$.

We now provide a sufficient condition on the subdifferential that ensures the energy is regular and the metric slope is a strong upper gradient. This generalized Ambrosio, Gigli, and Savaré's result that λ-convex energies are regular [2, Lemma 10.1.3].

Proposition 8.2. *Suppose* $\mathsf{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ *is proper, lower semicontinuous, and satisfies* $D(|\partial \mathsf{F}|) \subseteq \mathcal{P}_{2,ac}(\mathbb{R}^d)$. *Furthermore, suppose that there exists a continuous function* $\psi : [0, +\infty) \to [0, +\infty)$ *with* $\psi(0) = 0$ *so that, for any* $\xi \in \partial \mathsf{F}(\mu)$,

$$(8.1) \qquad \mathsf{F}(\nu) - \mathsf{F}(\mu) \geqslant \int \langle \xi, t_\mu^\nu - \mathrm{Id} \rangle \, d\mu - f(\mu, \nu) \psi(d_W(\mu, \nu)) d_W(\mu, \nu) \quad \text{for all } \nu \in D(\mathsf{F}),$$

where $f(\mu, \nu) = C(1 + \mathsf{F}(\mu) + \mathsf{F}(\nu))$ *for some* $C = C(d_W(\mu, \nu)) > 0$ *which is an increasing function of the distance from* μ *to* ν. *Then* F *is regular and the metric slope* $|\partial \mathsf{F}|$ *is a strong upper gradient for* F.

Proof. We begin by showing that F is regular. Consider a 2-Wasserstein convergent sequence $\mu_n \to \mu$ with strong subdifferentials $\xi_n \in \partial \mathsf{F}(\mu_n)$ satisfying criteria (i)-(iii) from Definition 8.1. First, we show that $\xi \in \partial \mathsf{F}(\mu)$. By assumption,

$$(8.2) \quad \mathsf{F}(\nu) - \mathsf{F}(\mu_n)$$
$$\geqslant \int \langle \xi_n, t_{\mu_n}^\nu - \mathrm{Id} \rangle \, d\mu_n - f(\mu_n, \nu) \psi(d_W(\mu_n, \nu)) d_W(\mu_n, \nu) \quad \text{for all } \nu \in \mathcal{P}_2(\mathbb{R}^d).$$

By the lower semicontinuity of F, we have $\liminf_{n \to +\infty} \mathsf{F}(\mu_n) \geqslant \mathsf{F}(\mu)$. Since $\sup_n \mathsf{F}(\mu_n) < +\infty$ and f is locally bounded on sublevels of F, we have $\sup_n f(\mu_n, \nu) < +\infty$. Since ψ is continuous,

$$\lim_{n \to +\infty} \psi(d_W(\mu_n, \nu)) d_W(\mu_n, \nu) = \psi(d_W(\mu, \nu)) d_W(\mu, \nu).$$

Furthermore, arguing as in [2, Lemma 10.1.3], we have

$$\lim_{n\to+\infty}\int\langle\xi_n,t^\nu_{\mu_n}-\mathrm{Id}\rangle\,d\mu_n=\int\langle\xi,t^\nu_\mu-\mathrm{Id}\rangle\,d\mu.$$

Therefore, sending $n\to+\infty$ in (8.2), we conclude that $\xi\in\partial\mathsf{F}(\mu)$.

Now, we show $\lim_{n\to+\infty}\mathsf{F}(\mu_n)=\mathsf{F}(\mu)$. Taking $\nu=\mu$ in (8.2) and sending $n\to+\infty$, the previous argument shows that the right hand side converges to 0. Thus,

$$\liminf_{n\to+\infty}\mathsf{F}(\mu)-\mathsf{F}(\mu_n)\geqslant 0\iff\liminf_{n\to+\infty}-\mathsf{F}(\mu_n)\geqslant-\mathsf{F}(\mu)\iff\limsup_{n\to+\infty}\mathsf{F}(\mu_n)\leqslant\mathsf{F}(\mu).$$

Combining this with the lower semicontinuity of F, we obtain that $\lim_{n\to+\infty}\mathsf{F}(\mu_n)=\mathsf{F}(\mu)$. Therefore, F is regular.

We now show that the metric slope $|\partial\mathsf{F}|$ is a strong upper gradient for F. We argue as in [2, Corollary 2.4.10]. Consider $\mu:(0,T)\to\mathcal{P}_2(\mathbb{R}^d)$ that is *absolutely continuous* in time, i.e., there exists $m\in L^1(0,T)$ so that inequality (2.3) holds. Then its metric derivate $|\mu'|(t)$ is well defined. It suffices to show that if $\mu(t)$ satisfies $|\partial\mathsf{F}|(\mu)|\mu'|\in L^1(0,T)$, then $\mathsf{F}(\mu(t))$ is absolutely continuous in time. By [2, Lemma 10.1.5], we have $|\partial\mathsf{F}|(\mu)<+\infty$ if and only if there exists some $\xi\in\partial\mathsf{F}(\mu)$ so that $|\partial\mathsf{F}|(\mu)=\|\xi\|_{L^2(\rho)}$.

We begin by showing that $\mathsf{F}(\mu(t))|\mu'|(t)\in L^1(0,T)$. Without loss of generality, $\mathsf{F}(\mu(t_*))<+\infty$ for some $t_*\in(0,T)$, or else $\mathsf{F}(\mu(t))\equiv+\infty$ is constant, hence absolutely continuous. Furthermore, up to reparametrizing time, we may also assume

$$\sup_{t\in(0,T)}C\psi(d_W(\mu(t),\mu(t_*)))d_W(\mu(t),\mu(t_*))\leqslant\frac{1}{2},$$

for $C=C(d_W(\mu,\nu))>0$ as in the definition of f. Applying inequality (8.1) with $\mu=\mu(t)$ and $\nu=\mu(t_*)$, we conclude that there exists $C'>0$ so that

$$\mathsf{F}(\mu(t))\leqslant\mathsf{F}(\mu(t_*))+|\partial\mathsf{F}|(\mu(t))d_W(\mu(t),\mu(t_*))$$
$$+f(\mu(t),\mu(t_*))\psi(d_W(\mu(t),\mu(t_*)))d_W(\mu(t),\mu(t_*))$$
$$\leqslant\mathsf{F}(\mu(t_*))+C'|\partial\mathsf{F}|(\mu(t))+\frac{1}{2}\Big(1+\mathsf{F}(\mu(t))+\mathsf{F}(\mu(t_*))\Big)$$

Rearranging and multiplying by $|\mu'|(t)$, we conclude that $\mathsf{F}(\mu(t))|\mu'|(t)\in L^1(0,T)$.

We now show that $\mathsf{F}(\mu(t))$ is absolutely continuous in time. Consider the compact subset $\mathcal{S}:=\{\mu(t):t\in[0,T]\}\subset\mathcal{P}_2(\mathbb{R}^d)$ and recall that the global slope on this subset

$$I_\mathsf{F}(\mu):=\sup_{\nu\in\mathcal{S},\nu\neq\mu}\frac{(\mathsf{F}(\mu)-\mathsf{F}(\nu))^+}{d_W(\mu,\nu)}$$

is a strong upper gradient [2, Theorem 1.2.5]. By inequality (8.1),

(8.3) $$\frac{(\mathsf{F}(\mu)-\mathsf{F}(\nu))^+}{d_W(\mu,\nu)}\leqslant|\partial\mathsf{F}|(\mu)+C\Big(1+\mathsf{F}(\mu)+\mathsf{F}(\nu)\Big)\psi(d_W(\mu,\nu)).$$

Furthermore, we may assume without loss of generality that $\mathsf{F}(\nu)\leqslant\mathsf{F}(\mu)$. Therefore,

$$I_\mathsf{F}(\mu(t))|\mu'|(t)\leqslant|\partial\mathsf{F}|(\mu(t))|\mu'|(t)+C\Big(1+2\mathsf{F}(\mu(t))\Big)\psi(\mathrm{diam}\mathcal{S})|\mu'|(t).$$

Therefore, we conclude that $I_\mathsf{F}(\mu(t))|\mu'|(t)\in L^1(0,T)$. Arguing as in [2, Theorem 1.2.5], we conclude that $\mathsf{F}(\mu(t))$ is absolutely continuous. $\square$

Corollary 8.3. *Suppose K satisfies hypotheses (LSC), (GF1)–(GF3) and $m\geqslant m_0$. Then E_m is* regular *and the metric slope $|\partial\mathsf{E}_m|$ is a strong upper gradient for E_m.*

Proof. The result is an immediate consequence of Proposition 4.3, Proposition 4.6, and Proposition 8.2, where we appeal to Proposition 4.1 to ensure that the estimates

$$C(1 + \|\rho_0\|_m + \|\rho_1\|_m) \leqslant C_r \left(1 + \|\rho_0\|_m^{1+\frac{1}{r}} + \|\rho_1\|_m^{1+\frac{1}{r}} \right) \leqslant \tilde{C}_r (1 + \mathsf{E}_m(\rho_1) + \mathsf{E}_m(\rho_2))$$

hold. □

8.2. **Power-Law Interaction Potentials.** In this section, we prove Proposition 4.5 and Proposition 4.7.

Proof of Proposition 4.5. The result extends [42, Proposition 4.6], where (GF1), (GF2), and (GF3) generalize [42, Assumption 4.1]. Define $\mathbf{t}_\alpha := (1 - \alpha)\,\mathrm{Id} + \alpha \mathbf{t}_{\rho_0}^{\rho_1}$ and let $\rho_\alpha := \mathbf{t}_\alpha \# \rho_0$ to be the Wasserstein geodesic from ρ_0 to ρ_1. By convexity of the L^p-norm, defined by equation (4.7), along Wasserstein geodesics, $\|\rho_\alpha\|_m \leqslant \max\{\|\rho_0\|_m, \|\rho_1\|_m\}$ for all $\alpha \in [0, 1]$. Then,

$$\begin{aligned}
\frac{d}{d\alpha} \mathsf{K}(\rho_\alpha) &= \lim_{h \to 0} \frac{1}{h}[\mathsf{K}(\rho_{\alpha+h}) - \mathsf{K}(\rho_\alpha)] \\
&= \lim_{h \to 0} \frac{1}{2h} \left[\int K * \rho_\alpha \, d\rho_{\alpha+h} - \int K * \rho_\alpha \, d\rho_\alpha \right] \\
&\quad + \frac{1}{2h} \left[\int K * \rho_{\alpha+h} \, d\rho_{\alpha+h} - \int K * \rho_{\alpha+h} \, d\rho_\alpha \right] \\
&= \lim_{h \to 0} \frac{1}{2h} \int [(K * \rho_\alpha) \circ \mathbf{t}_{\alpha+h} - (K * \rho_\alpha) \circ \mathbf{t}_\alpha] \, d\rho_0 \\
&\quad + \frac{1}{2h} \int [(K * \rho_{\alpha+h}) \circ \mathbf{t}_{\alpha+h} - (K * \rho_{\alpha+h}) \circ \mathbf{t}_\alpha] \, d\rho_0 \\
&= \lim_{h \to 0} \frac{1}{2h} \int \left(k_\alpha \circ \mathbf{t}_{\alpha+h} - k_\alpha \circ \mathbf{t}_\alpha \right) d\rho_0 + \frac{1}{2h} \int \left(k_{\alpha+h} \circ \mathbf{t}_{\alpha+h} - k_{\alpha+h} \circ \mathbf{t}_\alpha \right) d\rho_0,
\end{aligned}$$

for $k_\beta(x) := (K * \rho_\beta)(x)$. We consider both terms simultaneously by taking $\beta = \alpha$ or $\alpha + h$. By hypothesis (GF2), $k_\beta(x)$ is continuously differentiable with respect to x, so

$$k_\beta \circ \mathbf{t}_{\alpha+h} - k_\beta \circ \mathbf{t}_\alpha = \int_\alpha^{\alpha+h} \frac{d}{d\gamma} k_\beta \circ \mathbf{t}_\gamma \, d\gamma = \int_\alpha^{\alpha+h} \left\langle (\nabla K * \rho_\beta) \circ \mathbf{t}_\gamma, \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right\rangle d\rho_0 d\gamma.$$

Furthermore, since $\|\rho_\beta\|_m \leqslant \max\{\|\rho_0\|_m, \|\rho_1\|_m\}$, (GF1) ensures $\|\nabla K * \rho_\beta\|_{L^2(\rho_\gamma)} < +\infty$. Consequently, we may interchange the order of integration, and add and subtract to obtain

$$\begin{aligned}
(8.4) \quad \frac{d}{d\alpha} \mathsf{K}(\rho_\alpha) &= \lim_{h \to 0} \frac{1}{h} \int_\alpha^{\alpha+h} \int \left\langle (\nabla K * \rho_\alpha) \circ \mathbf{t}_\gamma, \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right\rangle d\rho_0 d\gamma \\
&\quad + \frac{1}{2h} \int_\alpha^{\alpha+h} \int \left\langle (\nabla K * \rho_{\alpha+h}) \circ \mathbf{t}_\gamma - (\nabla K * \rho_\alpha) \circ \mathbf{t}_\gamma, \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right\rangle d\rho_0 d\gamma.
\end{aligned}$$

In order to compute the limits on the right hand side, note that for any α, $\tilde{\alpha}$, β, $\tilde{\beta} \in [0, 1]$,

$$\int \left| (\nabla K * \rho_{\tilde{\alpha}}) \circ \mathbf{t}_{\tilde{\beta}} - (\nabla K * \rho_{\alpha}) \circ \mathbf{t}_{\beta} \right| \left| \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right| d\rho_0$$

$$\leqslant \Big[\| (\nabla K * \rho_{\tilde{\alpha}}) \circ \mathbf{t}_{\tilde{\beta}} - (\nabla K * \rho_{\tilde{\alpha}}) \circ \mathbf{t}_{\beta} \|_{L^2(\rho_0)}$$

$$+ \| \nabla K * \rho_{\tilde{\alpha}} - \nabla K * \rho_{\alpha} \|_{L^2(\rho_\beta)} \Big] \| \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \|_{L^2(\rho_0)}$$

(8.5)
$$\leqslant f(\rho_0, \rho_1) \left[\left\| \psi \big(|\mathbf{t}_{\tilde{\beta}} - \mathbf{t}_{\beta}|^2 \big) \right\|_{L^1(\rho_0)}^{1/2} + \psi(d_W(\rho_{\tilde{\alpha}}, \rho_{\alpha})) \right] d_W(\rho_0, \rho_1)$$

$$\leqslant f(\rho_0, \rho_1) \left[\psi \left(\| \mathbf{t}_{\tilde{\beta}} - \mathbf{t}_{\beta} \|_{L^2(\rho_0)}^2 \right)^{1/2} + \psi \left(d_W(\rho_{\tilde{\alpha}}, \rho_{\alpha}) \right) \right] d_W(\rho_0, \rho_1)$$

$$\leqslant f(\rho_0, \rho_1) \left[\psi \left(|\tilde{\beta} - \beta|^2 d_W(\rho_0, \rho_1)^2 \right)^{1/2} + \psi \left(|\tilde{\alpha} - \alpha| d_W(\rho_0, \rho_1) \right) \right] d_W(\rho_0, \rho_1).$$

In the second inequality, we use hypotheses (GF2) and (GF3) and the fact that the 2-Wasserstein distance dominates the $(2 - \epsilon)$-Wasserstein distance for all $\epsilon > 0$; see inequality 2.2. In the third inequality we use Jensen's inequality for the concave function $\psi(s)$.

When $\alpha = \tilde{\alpha}$, this estimate ensures

$$\beta \mapsto \int \left\langle (\nabla K * \rho_{\alpha}) \circ \mathbf{t}_{\beta}, \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right\rangle d\rho_0$$

is continuous, so that the first term in (8.4) converges to $\int \left\langle (\nabla K * \rho_{\alpha}) \circ \mathbf{t}_{\alpha}, \mathbf{t}_{\rho_0}^{\mu_1} - \mathrm{Id} \right\rangle d\rho_0$. Likewise, when $\beta = \tilde{\beta}$, this estimate guarantees that the second term in (8.4) is bounded by

$$\lim_{h \to 0} \frac{f(\rho_0, \rho_1)}{2h} \int_{\alpha}^{\alpha+h} \psi \Big(h \, d_W(\rho_0, \rho_1) \Big) d_W(\rho_0, \rho_1) \, d\gamma = 0.$$

Therefore, we conclude

$$\frac{d}{d\alpha} \mathsf{K}(\rho_{\alpha}) = \int \left\langle (\nabla K * \rho_{\alpha}) \circ \mathbf{t}_{\alpha}, \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right\rangle d\rho_0.$$

By (8.5) again, $\frac{d}{d\alpha} \mathsf{K}(\mu_{\alpha})$ is continuous for $\alpha \in [0, 1]$. Therefore,

(8.6)
$$\mathsf{K}(\rho_1) = \mathsf{K}(\rho_0) + \int_0^1 \frac{d}{d\alpha} \mathsf{K}(\rho_{\alpha}) \, d\alpha$$

$$= \mathsf{K}(\rho_0) + \frac{d}{d\alpha} \mathsf{K}(\rho_{\alpha}) \Big|_{\alpha=0} + \int_0^1 \int \left\langle (\nabla K * \rho_{\alpha}) \circ \mathbf{t}_{\alpha} - \nabla K * \rho_0, \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right\rangle d\rho_0 d\alpha.$$

To prove the result, it suffices to show that the third term is $o(d_W(\rho_0, \rho_1))$. This follows by a final application of inequality (8.5):

$$\left| \int_0^1 \int \left\langle (\nabla K * \rho_{\alpha}) \circ \mathbf{t}_{\alpha} - \nabla K * \rho_0, \mathbf{t}_{\rho_0}^{\rho_1} - \mathrm{Id} \right\rangle d\rho_0 d\alpha \right|$$

$$\leqslant f(\rho_0, \rho_1) \left[\psi \left(d_W(\rho_0, \rho_1)^2 \right)^{1/2} + \psi \Big(d_W(\rho_0, \rho_1) \Big) \right] d_W(\rho_0, \rho_1)$$

Finally, by Lemma 8.4, there exists c, which is an increasing function of $d_W(\rho_0, \rho_1)$, so that $\sqrt{\psi(s^2)} \leqslant c\psi(s)$. Therefore, up to increasing the constant in the definition of $f(\rho_0, \rho_1)$, inequality (8.7) is bounded by

$$f(\rho_0, \rho_1) \, \psi(d_W(\rho_0, \rho_1)) \, d_W(\rho_0, \rho_1),$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

We now turn to the proof of Proposition 4.7.

Proof of Proposition 4.7. Our proof follows a similar approach as [2, Theorem 10.4.13], generalizing this result to nonconvex, singular interaction potentials K. We begin by proving (4.9). Note that the implication $\Longleftarrow$ is immediate, as the definition of the subdifferential $\xi \in \partial \mathsf{E}_m(\rho)$ requires $\|\xi\|_{L^2(\rho)} < +\infty$. (See inequality (2.7).)

Suppose $|\partial \mathsf{E}_m|(\rho) < +\infty$. Since $\rho \in D(\mathsf{E}_m)$, we have $\|\rho\|_m < +\infty$. Fix $\xi \in L^2(\rho)$ and define $r_\alpha = (1-\alpha)\operatorname{Id} +\alpha\xi$ and $\rho_\alpha := (r_\alpha)_\# \rho$. Suppose either (i) $\xi - \operatorname{Id} \in C_c^\infty(\mathbb{R}^d;\mathbb{R}^d)$ or (ii) $\xi = 0$. Note that both assumptions ensure that ξ is differentiable almost everywhere (cf. Aleksandrov's theorem, [2, Theorem 5.5.4]) and there exists $\alpha_0, C > 0$ so that for all $\alpha \in [0, \alpha_0]$, $\|\rho_\alpha\|_m < C$(cf. [2, Lemma 5.5.3]).

Under either assumption (i) or (ii), the definition of the metric slope (2.5), Proposition 4.5 for K, and [2, Lemma 10.4.4] for S_m, ensure

$$
\begin{aligned}
|\partial \mathsf{E}_m(\rho)| \, \|\xi - \operatorname{Id}\|_{L^2(\rho)} &\geqslant \limsup_{\alpha \to 0} \frac{\mathsf{E}_m(\rho_\alpha) - \mathsf{E}_m(\rho)}{d_W(\rho, \rho_\alpha)} \|\xi - \operatorname{Id}\|_{L^2(\rho)} \\
&= \limsup_{\alpha \to 0} \frac{\mathsf{K}(\rho_\alpha) - \mathsf{K}(\rho)}{\alpha} + \frac{\mathsf{S}_m(\rho_\alpha) - \mathsf{S}_m(\rho)}{\alpha} \\
&= \int \langle \nabla K * \rho, \xi - \operatorname{Id} \rangle \, d\rho - \int \rho^m \nabla \cdot (\xi - \operatorname{Id}) \, dx.
\end{aligned}
$$
(8.7)

Thus, by Hölder's inequality and hypothesis (GF1), there exists $C' > 0$ depending on ρ so that

$$
(8.8) \qquad |\partial \mathsf{E}_m(\rho)| \, \|\xi - \operatorname{Id}\|_{L^2(\rho)} + C'\|\xi - \operatorname{Id}\|_{L^2(\rho)} \geqslant -\int \rho^m \nabla \cdot (\xi - \operatorname{Id}) \, dx.
$$

First, we suppose ξ satisfies assumption (i). As inequality (8.8) holds for all $f = \xi - \operatorname{Id} \in C_c^\infty(\mathbb{R}^d)$, this implies ρ^m is a function of bounded variation and the right hand side may be rewritten as $\int \nabla \rho^m \cdot (\xi - \operatorname{Id}) \, dx$. Applying inequality (8.8) again, we obtain $\nabla \rho^m \in L^2(\rho)$. Returning back to (8.7), we obtain that

$$
(8.9) \qquad \left\| (\nabla K * \rho) + \frac{\nabla \rho^m}{\rho} \right\|_{L^2(\rho)} \leqslant |\partial \mathsf{E}_m|(\rho).
$$

Next, suppose $\xi = 0$. Then inequality (8.8) gives

$$
(|\partial \mathsf{E}_m(\rho)| + C')M_2(\rho)^{1/2} \geqslant d\left(\int \rho^m \, dx \right).
$$

hence $\rho^m \in W^{1,1}(\mathbb{R}^d)$.

To complete the proof, it suffices to show $(\nabla K * \rho) + \frac{\nabla \rho^m}{\rho} \in \partial \mathsf{E}_m(\rho)$. Remark 2.1 ensures equality must hold in (8.9). Since the subdifferential $\partial \mathsf{E}_m(\rho)$ is a convex subset of $L^2(\rho)$ and the $L^2(\rho)$-norm is strictly convex, the element of $\partial \mathsf{E}_m(\rho)$ with minimal $L^2(\rho)$-norm is unique.

By Proposition 4.6, to show $(\nabla K * \rho) + \frac{\nabla \rho^m}{\rho} \in \partial \mathsf{E}_m(\rho)$ it suffices to show that

$$
\begin{aligned}
\mathsf{E}_m(\nu) &- \mathsf{E}_m(\rho) \\
&\geqslant \int \left\langle (\nabla K * \rho) + \frac{\nabla \rho^m}{\rho}, \mathbf{t}_\rho^\nu - \operatorname{Id} \right\rangle d\rho - f(\rho, \nu)\psi(d_W(\rho, \nu))d_W(\rho, \nu), \quad \forall \, \nu \in D(\mathsf{E}_m).
\end{aligned}
$$

Since $\rho, \nu \in D(\mathsf{E}_m)$, by Proposition 4.5, for the subdifferential of K, and by [2, Theorem 10.4.6], for the subdifferential of S_m,

$$(8.10) \qquad \mathsf{K}(\nu) - \mathsf{K}(\rho) = \int \left\langle \nabla K * \rho, \mathbf{t}_\rho^\nu - \mathrm{Id} \right\rangle d\rho - f(\rho, \nu)\psi(d_W(\rho, \nu))d_W(\rho, \nu),$$

$$(8.11) \qquad \mathsf{S}_m(\nu) - \mathsf{S}_m(\rho) \geqslant \int \left\langle \frac{\nabla \rho^m}{\rho}, \mathbf{t}_\rho^\nu - \mathrm{Id} \right\rangle d\rho$$

Adding together inequalities (8.10) and (8.11) gives the result. $\qquad\square$

8.3. Elementary Bounds.

Lemma 8.4. *Suppose $\psi : [0, +\infty) \to [0, +\infty)$ is a continuous, nondecreasing, concave function with $\psi(0) = 0$. Then for any $\bar{s}$, there exists $\bar{C}$ so that*

$$\psi(s^2) \leqslant \bar{C}\psi(s)^2, \qquad \text{for all } s \in [0, \bar{s}].$$

Proof. Since $\psi(s)$ is concave and $\psi(0) = 0$, $\psi(s)/s$ is a decreasing function and $\partial^+\psi(0) = \lim_{s\to 0^+} \psi(s)/s$ exists. First, suppose $\partial^+\psi(0) = 0$. Since ψ is concave and nondecreasing, this implies $\psi \equiv 0$, and the result holds.

Now, suppose $\partial^+\psi(0) > 0$. It suffices to show that $\psi(s)^2/\psi(s^2)$ is uniformly bounded below on $[0, \bar{s}]$. Since $\partial^+\psi(0) > 0$ and ψ is increasing, $\psi(s) > 0$ for all $s > 0$ and $\psi(s)^2/\psi(s^2)$ is continuous and positive on $(0, \bar{s}]$, hence bounded below away from $s = 0$. Furthermore,

$$\lim_{s\to 0} \frac{\psi(s)^2}{\psi(s^2)} = \partial^+\psi(0) > 0.$$

Thus $\psi(s)^2/\psi(s^2)$ is bounded below on $[0, \bar{s}]$. $\qquad\square$

Lemma 8.5. *For $s \geqslant 0$ and $0 < b < 1/2$, define $\psi(s, a) := s^{(1-a)}$. Then,*

$$|\psi(s, b) - \psi(s, 0)| \leqslant b(1 + s^{2-b}).$$

Proof. Define $\psi(s, a) := s^{(1-a)}$, so $\frac{d}{da}\psi(s, a) = -s^{1-a}\log(s)$. By the mean value theorem, for $0 < b < 1/2$, there exists $a \in [0, b]$ so that

$$\psi(s, b) - \psi(s, 0) = b\frac{d}{da}\psi(s, a).$$

If $s \in [0, 1]$, $\left|\frac{d}{da}\psi(s, a)\right| \leqslant 1$, so

$$(8.12) \qquad\qquad\qquad |\psi(s, b) - \psi(s, 0)| \leqslant b.$$

If $s \geqslant 1$, we claim that

$$(8.13) \qquad\qquad\qquad |\psi(s, b) - \psi(s, 0)| \leqslant bs^{2-b}$$

This holds since, for $s > 1$,

$$|\psi(s, b) - \psi(s, 0)| = \left|s^{1-b} - s\right| = s - s^{1-b} \leqslant bs^{2-b} \iff s^{1+b} - s \leqslant bs^2 \iff s^{1+b} \leqslant bs^2 + s.$$

This is true at $s = 1$, so it suffices to show the derivative with respect to s of the right-hand side is larger than the derivative of the left-hand side, i.e.,

$$(1 + b)s^b \leqslant 2bs + 1$$

This is also true at $s = 1$, so differentiating again, it suffices to show

$$(1 + b)bs^{(b-1)} \leqslant 2b$$

This holds since $s \geqslant 1$ and $0 < b^2 < b < 1/2$. Combining (8.12) and (8.13) gives the result. $\quad\square$

References

[1] D. Alexander, I. Kim, and Y. Yao, "Quasi-static evolution and congested crowd transport," *Nonlinearity*, vol. 27, no. 4, pp. 823–858, 2014.

[2] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., ser. Lectures in Mathematics ETH Zürich. Basel: Birkhäuser Verlag, 2008.

[3] L. Ambrosio, E. Mainini, and S. Serfaty, "Gradient flow of the Chapman-Rubinstein-Schatzman model for signed vortices," *Ann. Inst. H. Poincaré Anal. Non Linéaire*, vol. 28, no. 2, pp. 217–246, 2011.

[4] L. Ambrosio and S. Serfaty, "A gradient flow approach to an evolution problem arising in superconductivity," *Comm. Pure Appl. Math.*, vol. 61, no. 11, pp. 1495–1539, 2008.

[5] L. Ambrosio and A. Pratelli, "Existence and stability results in the l 1 theory of optimal transportation," in *Optimal transportation and applications*. Springer, 2003, pp. 123–160.

[6] D. Balagué, J. Carrillo, and Y. Yao, "Confinement for repulsive-attractive kernels," *Discrete Contin. Dyn. Syst. Ser. B*, vol. 19, no. 5, pp. 1227–1248, 2014.

[7] M. Bandegi and D. Shirokoff, "Approximate global minimizers to pairwise interaction problems via convex relaxation," *SIAM Journal on Applied Dynamical Systems*, vol. 17, no. 1, pp. 417–456, 2018.

[8] J. Bedrossian, "Global minimizers for free energies of subcritical aggregation equations with degenerate diffusion," *Appl. Math. Lett.*, vol. 24, no. 11, pp. 1927–1932, 2011.

[9] ——, "Intermediate asymptotics for critical and supercritical aggregation equations and Patlak-Keller-Segel models," *Commun. Math. Sci.*, vol. 9, no. 4, pp. 1143–1161, 2011.

[10] J. Bedrossian, N. Rodríguez, and A. Bertozzi, "Local and global well-posedness for aggregation equations and Patlak-Keller-Segel models with degenerate diffusion," *Nonlinearity*, vol. 24, no. 6, pp. 1683–1714, 2011.

[11] J. Bedrossian and N. Rodríguez, "Inhomogeneous Patlak-Keller-Segel models and aggregation equations with nonlinear diffusion in $\mathbb{R}^d$," *Discrete Contin. Dyn. Syst. Ser. B*, vol. 19, no. 5, pp. 1279–1309, 2014.

[12] A. Bertozzi and D. Slepčev, "Existence and uniqueness of solutions to an aggregation equation with degenerate diffusion," *Commun. Pure Appl. Anal.*, vol. 9, no. 6, pp. 1617–1637, 2010.

[13] A. Bertozzi, H. Sun, T. Kolokolnikov, D. Uminsky, and J. Von Brecht, "Ring patterns and their bifurcations in a nonlocal model of biological swarms," *Comm. Math. Sci.*, vol. 13, no. 4, pp. 955–985, 2015.

[14] S. Bian and J. G. Liu, "Dynamic and steady states for multi-dimensional Keller-Segel model with diffusion exponent $m > 0$," *Comm. Math. Phys.*, vol. 323, pp. 1017–1070, 2013.

[15] A. Blanchet, E. Carlen, and J. Carrillo, "Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model," *J. Funct. Anal.*, vol. 262, no. 5, pp. 2142–2230, 2012.

[16] A. Blanchet, J. Carrillo, and P. Laurençot, "Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions," *Calc. Var. Partial Differential Equations*, vol. 35, no. 2, pp. 133–168, 2009.

[17] A. Blanchet, V. Calvez, and J. A. Carrillo, "Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model," *SIAM J. Numer. Anal.*, vol. 46, no. 2, pp. 691–721, 2008.

[18] A. Blanchet, J. A. Carrillo, and N. Masmoudi, "Infinite time aggregation for the critical Patlak-Keller-Segel model in $\mathbb{R}^2$," *Comm. Pure Appl. Math.*, vol. 61, no. 10, pp. 1449–1481, 2008.

[19] A. Blanchet, J. Dolbeault, and B. Perthame, "Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions," *Electron. J. Differential Equations*, no. 44, 2006.

[20] A. Burchard, R. Choksi, and I. Topaloglu, "Nonlocal shape optimization via interactions of attractive and repulsive potentials," *Indiana Univ. Math. J.*, vol. 67, no. 1, pp. 375–395, 2018.

[21] L. A. Caffarelli and A. Friedman, "Asymptotic behavior of solutions of $u_t = \Delta u^m$ as $m \to \infty$," *Indiana Univ. Math. J.*, vol. 36, no. 4, pp. 711–728, 1987.

[22] V. Calvez and J. A. Carrillo, "Refined asymptotics for the subcritical Keller-Segel system and related functional inequalities," *Proc. Amer. Math. Soc.*, vol. 140, no. 10, pp. 3515–3530, 2012.

[23] V. Calvez, J. A. Carrillo, and F. Hoffmann, "Equilibria of homogeneous functionals in the fair-competition regime," *Nonlinear Analysis*, vol. 159, pp. 85–128, 2017.

[24] ——, "The geometry of diffusing and self-attracting particles in a one-dimensional fair-competition regime," in *Nonlocal and Nonlinear Diffusions and Interactions: New Methods and Directions*. Springer, 2017, pp. 1–71.

[25] G. Carlier, V. Duval, G. Peyré, and B. Schmitzer, "Convergence of entropic schemes for optimal transport and gradient flows," *SIAM Journal on Mathematical Analysis*, vol. 49, no. 2, pp. 1385–1418, 2017.

[26] J. Carrillo, S. Hittmeir, B. Volzone, and Y. Yao, "Nonlinear aggregation-diffusion equations: Radial symmetry and long time asymptotics," *arXiv preprint arXiv:1603.07767*, 2016.

[27] J. Carrillo, S. Lisini, and E. Mainini, "Uniqueness for Keller-Segel-type chemotaxis models," *Discrete Contin. Dyn. Syst.*, vol. 34, no. 4, pp. 1319–1338, 2014.

[28] J. Carrillo, R. McCann, and C. Villani, "Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates," *Rev. Mat. Iberoamericana*, vol. 19, no. 3, pp. 971–1018, 2003.

[29] ——, "Contractions in the 2-Wasserstein length space and thermalization of granular media," *Arch. Ration. Mech. Anal.*, vol. 179, no. 2, pp. 217–263, 2006.

[30] J. A. Carrillo, K. Craig, and Y. Yao, "Aggregation-diffusion equations: dynamics, asymptotics, and singular limits," *arXiv preprint arXiv:1810.03634*, 2018.

[31] J. A. Carrillo, F. Hoffmann, E. Mainini, and B. Volzone, "Ground states in the diffusion-dominated regime," *Calc. Var. Partial Differential Equations*, vol. 57, no. 5, pp. Art. 127, 28, 2018.

[32] J. A. Carrillo and Y. Sugiyama, "Compactly supported stationary states of the degenerate Keller-Segel system in the diffusion-dominated regime," *Indiana Univ. Math. J.*, vol. 67, no. 6, pp. 2279–2312, 2018.

[33] J. A. Carrillo and J. Wang, "Uniform in time L^∞-estimates for nonlinear aggregation-diffusion equations," *arXiv preprint arXiv:1712.09541*, 2017.

[34] J. A. Carrillo, D. Castorina, and B. Volzone, "Ground States for Diffusion Dominated Free Energies with Logarithmic Interaction," *SIAM Journal on Mathematical Analysis*, vol. 47, no. 1, pp. 1–25, Jan. 2015.

[35] J. A. Carrillo, K. Craig, and F. S. Patacchini, "A blob method for diffusion," *Calc. Var. Partial Differential Equations*, vol. 58, no. 2, pp. Art. 53, 53, 2019.

[36] J. A. Carrillo, K. Craig, L. Wang, and C. Wei, "A primal dual method for Wasserstein gradient flows," *in preparation*.

[37] L. Chayes, I. Kim, and Y. Yao, "An aggregation equation with degenerate diffusion: Qualitative property of solutions," *SIAM Journal on Mathematical Analysis*, vol. 45, no. 5, pp. 2995–3018, 2013.

[38] L. Chizat and S. Di Marino, "A tumor growth model of Hele-Shaw type as a gradient flow," *arXiv preprint arXiv:1712.06124*, 2017.

[39] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard, "Scaling algorithms for unbalanced optimal transport problems," *Math. Comp.*, vol. 87, no. 314, pp. 2563–2609, 2018.

[40] R. Choksi, R. C. Fetecau, and I. Topaloglu, "On minimizers of interaction functionals with competing attractive and repulsive potentials," *Ann. Inst. H. Poincaré Anal. Non Linéaire*, vol. 32, no. 6, pp. 1283–1305, 2015.

[41] R. Choksi, C. B. Muratov, and I. Topaloglu, "An old problem resurfaces nonlocally: Gamow's liquid drops inspire today's research and applications," *Notices Amer. Math. Soc.*, vol. 64, no. 11, pp. 1275–1283, 2017.

[42] K. Craig, "Nonconvex gradient flow in the wasserstein metric and applications to constrained nonlocal interactions," *Proceedings of the London Mathematical Society*, vol. 114, no. 1, pp. 60–102, 2017.

[43] K. Craig and A. L. Bertozzi, "A blob method for the aggregation equation," *Math. Comp.*, vol. 85, no. 300, pp. 1681–1717, 2016.

[44] K. Craig, I. Kim, and Y. Yao, "Congested aggregation via Newtonian interaction," *Arch. Ration. Mech. Anal.*, vol. 227, no. 1, pp. 1–67, 2018.

[45] K. Craig and I. Topaloglu, "Convergence of regularized nonlocal interaction energies," *SIAM J. Math. Anal.*, vol. 48, no. 1, pp. 34–60, 2016.

[46] R. L. Frank and E. H. Lieb, "A "liquid-solid" phase transition in a simple model for swarming, based on the "no flat-spots" theorem for subharmonic functions," *Indiana Univ. Math. J.*, vol. 67, no. 4, pp. 1547–1569, 2018.

[47] T. Gallouët, M. Laborde, and L. Monsaingeon, "An unbalanced optimal transport splitting scheme for general advection-reaction-diffusion problems," *ESAIM: COCV*, 2019, to appear.

[48] I. Kim and N. Požár, "Porous medium equation to Hele-Shaw flow with general initial density," *Trans. Amer. Math. Soc.*, vol. 370, no. 2, pp. 873–909, 2018.

[49] I. Kim and Y. Yao, "The Patlak-Keller-Segel model and its variations: properties of solutions via maximum principle," *SIAM J. Math. Anal.*, vol. 44, no. 2, pp. 568–602, 2012.

[50] G. Leoni, *A first course in Sobolev spaces*, 2nd ed., ser. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2017, vol. 181.

[51] E. Lieb and M. Loss, *Analysis*, 2nd ed., ser. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2001, vol. 14.

[52] P.-L. Lions, "The concentration-compactness principle in the calculus of variations. The locally compact case. I," *Ann. Inst. H. Poincaré Anal. Non Linéaire*, vol. 1, no. 2, pp. 109–145, 1984.

[53] J.-G. Liu, L. Wang, and Z. Zhou, "Positivity-preserving and asymptotic preserving method for 2d Keller-Segel equations," *Mathematics of Computation*, vol. 87, no. 311, pp. 1165–1189, 2018.

[54] G. Loeper, "Uniqueness of the solution to the vlasov–poisson system with bounded density," *Journal de mathématiques pures et appliquées*, vol. 86, no. 1, pp. 68–79, 2006.

[55] O. Lopes, "Uniqueness and radial symmetry of minimizers for a nonlocal variational problem," *Commun. Pure Appl. Anal.*, vol. 18, no. 5, pp. 2265–2282, 2019.

[56] B. Maury, A. Roudneff-Chupin, and F. Santambrogio, "A macroscopic crowd motion model of gradient flow type," *Math. Models Methods Appl. Sci.*, vol. 20, no. 10, pp. 1787–1821, 2010.

[57] B. Maury, A. Roudneff-Chupin, F. Santambrogio, and J. Venel, "Handling congestion in crowd motion modeling," *Netw. Heterog. Media*, vol. 6, no. 3, pp. 485–519, 2011.

[58] R. McCann, "A convexity principle for interacting gases," *Adv. Math.*, vol. 128, no. 1, pp. 153–179, 1997.

[59] A. Mellet, B. Perthame, and F. Quiros, "A Hele–Shaw problem for tumor growth," *Journal of Functional Analysis*, vol. 273, no. 10, pp. 3061–3093, 2017.

[60] Y. Mizuta, *Potential theory in Euclidean spaces.* Gakkōtosho Tokyo, 1996.

[61] B. Perthame, F. Quirós, and J. L. Vázquez, "The Hele–Shaw asymptotics for mechanical models of tumor growth," *Archive for Rational Mechanics and Analysis*, vol. 212, no. 1, pp. 93–127, 2014.

[62] G. Peyré, "Entropic approximation of wasserstein gradient flows," *SIAM Journal on Imaging Sciences*, vol. 8, no. 4, pp. 2323–2351, 2015.

[63] G. Rein, "Reduction and a concentration-compactness principle for energy-Casimir functionals," *SIAM J. Math. Anal.*, vol. 33, no. 4, pp. 896–912, 2001.

[64] S. Serfaty, "Gamma-convergence of gradient flows on Hilbert and metric spaces and applications," *Discrete Contin. Dyn. Syst.*, vol. 31, no. 4, pp. 1427–1451, 2011.

[65] R. Simione, D. Slepčev, and I. Topaloglu, "Existence of ground states of nonlocal-interaction energies," *J. Stat. Phys.*, vol. 159, no. 4, pp. 972–986, 2015.

[66] E. M. Stein, *Singular integrals and differentiability properties of functions (PMS-30).* Princeton university press, 2016, vol. 30.

[67] Y. Sugiyama *et al.*, "Global existence in sub-critical cases and finite time blow-up in super-critical cases to degenerate Keller-Segel systems," *Differential and Integral Equations*, vol. 19, no. 8, pp. 841–876, 2006.

[68] A. van der Vaart and J. Wellner, *Weak Convergence and Empirical Processes: With Applications to Statistics*, ser. Springer Series in Statistics. Springer, 1996.

[69] C. Villani, *Topics in optimal transportation*, ser. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2003, vol. 58.

[70] Y. Yao and A. Bertozzi, "Blow-up dynamics for the aggregation equation with degenerate diffusion," *Phys. D*, vol. 260, pp. 77–89, 2013.

[71] Y. Yao, "Asymptotic behavior for critical Patlak-Keller-Segel model and a repulsive-attractive aggregation equation," *Ann. Inst. H. Poincaré Anal. Non Linéaire*, vol. 31, no. 1, pp. 81–101, 2014.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SANTA BARBARA, CA
E-mail address: `kcraig@math.ucsb.edu`

DEPARTMENT OF MATHEMATICS AND APPLIED MATHEMATICS, VIRGINIA COMMONWEALTH UNIVERSITY, RICHMOND, VA
E-mail address: `iatopaloglu@vcu.edu`